5-2008

A Comparative Analysis of Hyperspectral Target Detection Algorithms in the Presence of Misregistered Data

Jason T. Casey

A Comparative Analysis of Hyperspectral Target Detection Algorithms in the Presence of Misregistered Data

by

Jason T. Casey

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science
in the Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

May, 2008

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Jason T. Casey
has been examined and approved by the
thesis committee as satisfactory for the
thesis required for the
M.S. degree in Imaging Science

Dr. John P. Kerekes, Thesis Advisor

Dr. Emmett J. Ientilucci

Dr. Harvey E. Rhody

17 July 2008
Date

A Comparative Analysis of Hyperspectral Target Detection Algorithms in the Presence of Misregistered Data

by

Jason T. Casey

B.A. Physics, SUNY Geneseo, 2005
B.A. Computer Science, SUNY Geneseo, 2005

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science
in the Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

May, 2008

Signature of the Author _____

Accepted by _____ 4 Sep 2008
Coordinator, M.S. Degree Program Date

THESIS RELEASE PERMISSION

ROCHESTER INSTITUTE OF TECHNOLOGY

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

Title of Thesis:

A Comparative Analysis of Hyperspectral Target Detection Algorithms in the Presence of Misregistered Data

I, Jason T. Casey, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature _____

Date

A Comparative Analysis of Hyperspectral Target Detection Algorithms in the Presence of Misregistered Data

by

Jason T. Casey

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Master of Science Degree
at the Rochester Institute of Technology

Abstract

Line scanning hyperspectral imaging systems are capable of capturing accurate spatial and spectral information about a scene. These data can be useful for detecting sub-pixel targets. Such systems, however, may be limited by certain key characteristics in their design. Systems employing multiple spectrometers, or that collect data from multiple focal planes may suffer an inherent misregistration between sets of collected spectral bands. In order to utilize the full spectrum for target detection purposes, the sets of bands must be registered to each other as precisely as possible. Perfect registration is not possible, due to both the sensor design, and variation in sensor orientation during data acquisition. The issue can cause degradation in the performance of various target detection algorithms. An analysis of algorithms is necessary to determine which perform well when working with misregistered data. In addition, new algorithms may need to be developed which are more robust in these conditions. The work set forth in this thesis will improve the registration between spectral bands in a line scanning hyperspectral sensor by using a geometric model of the sensor along with aircraft orientation parameters to pair sets of image pixels based on their ground locations. Synthetic scenes were created and band-to-band misregistration was induced between the VIS and NIR spectral channels to test the performance of various hyperspectral target detection algorithms when applied to misregistered hyperspectral data. The results for this case studied show geometric algorithms perform well using only the VIS portion of the EM spectrum, and do not always benefit from the addition of NIR bands, even for small amounts of misregistration. Stochastic algorithms appear to be more robust than geometric algorithms for datasets with band-to-band

misregistration. The stochastic algorithms tested often benefit from the addition of NIR bands, even for large amounts of misregistration.

Acknowledgements

I would like to thank my advisor, Dr. John Kerekes, for his advise, support and guidance throughout my research and the writing of this thesis. I would also like to thank my committee members, Dr. Harvey Rhody and Dr. Emmett Ientilucci for their help and support. Thanks to Donald McKeown for helping to develop the MISI geometric sensor model. Thanks as well, to Dr. David Messinger for hiring me as a research assistant in the Summer of 2005 and getting me interested in the Imaging Science program at RIT. Finally I'd like to thank any faculty, staff, and students at the Center for Imaging Science who have helped me during my time at RIT, there are too many of you to list. I greatly appreciate everything you've done to help.

This material is based on research sponsored by AFRL/SNAT under agreement number FA8650-04-1-1717 (BAA 4-03-SNK Amendment 3). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views expressed in this thesis are those of the author and do not reflect official policy or position of the United States Air Force, the Department of Defense, or the U.S. Government.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Remote sensing can be described as gathering information about objects without coming into physical contact with them. As humans we use our eyes to remotely sense things basically every moment we are awake. Remote sensing as most scientists know it, however, is usually thought of as gathering information about the earth by looking down at it from overhead. This is usually done using airborne or satellite sensors sensing electromagnetic radiation.

Remote sensing as we think of it began in the mid to late 1800's after techniques were developed to allow short exposure times for film. Cameras were mounted on balloons to take downward looking images from the cameras hanging hundreds of meters in the air. Years later in 1946, a V-2 missile would leave the earths atmosphere, and an onboard camera would capture the first images of our planet from space. These images were taken with a simple 35mm motion picture camera. (1) Since then, equipment has become much more sophisticated, and can give us much more information about the a scene.

The type of imaging system familiar to most people would probably be a framing camera, or a framing array as shown in Figure 1.1. Framing cameras operate in the same way as cameras used for traditional photography. A framing camera consists of 3 parts: (1) the camera magazine, which contains film-advancing, and film flattening mechanisms, (2) the camera body, which houses the mechanical driving mechanisms of the camera, and (3) the lens cone assembly, which contains the lens, shutter, diaphram, and any filters that might be used with the camera. Framing arrays basically operate in the same manner, but film has been replaced by some type of detector array which is sensitive to incident radiation. Framing cameras and arrays are useful for aerial mapping. Film cameras have been used extensively for remote sensing

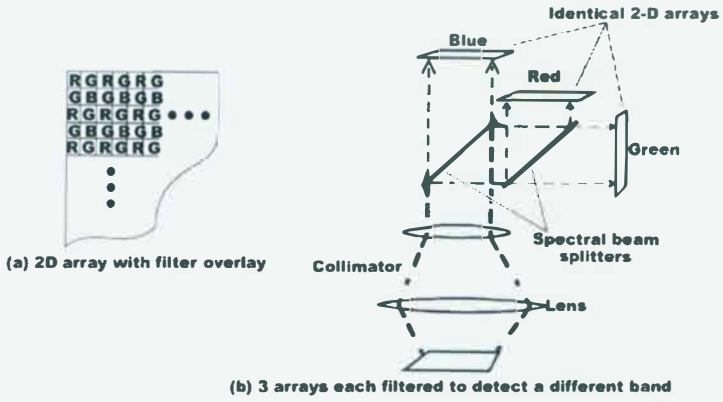


Figure 1.1: Figure of a framing system with three separate focal planes and framing arrays. (2)

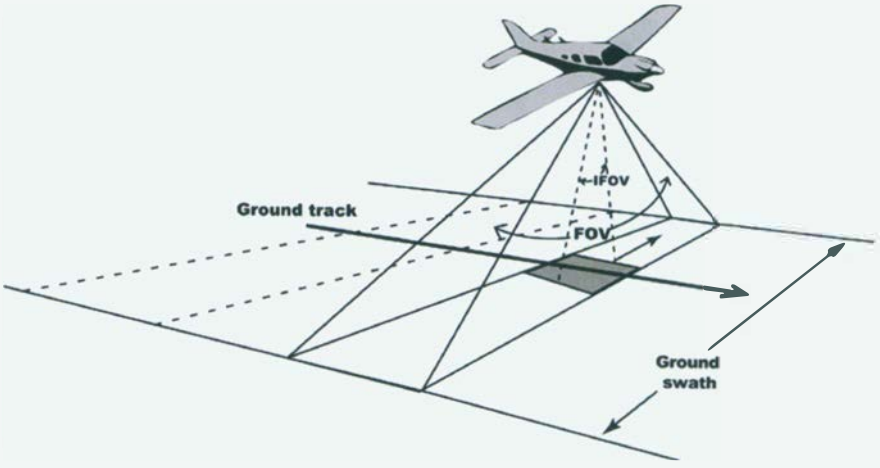


Figure 1.2: Concept of a line scanning system (2)

applications, especially when using infrared film, which is very useful for vegetative analysis. Useful spectral information can also be extracted using a framing array along with a set of filters, which transmit only light of certain wavelengths.

Hyperspectral sensors capture images at tens to hundreds of narrow, continuous spectral bands. There are a variety of hyperspectral sensor designs. Line scanning systems collect data one pixel at a time in the direction perpendicular to the flightpath of the aircraft. A demonstration of how a line scanning system collects imagery is given in Figure 1.2

A scanning mirror is used to change the location the sensor is looking at on the ground. Each pixel is recorded as the scan mirror rotates, which constructs a line of data. As the aircraft

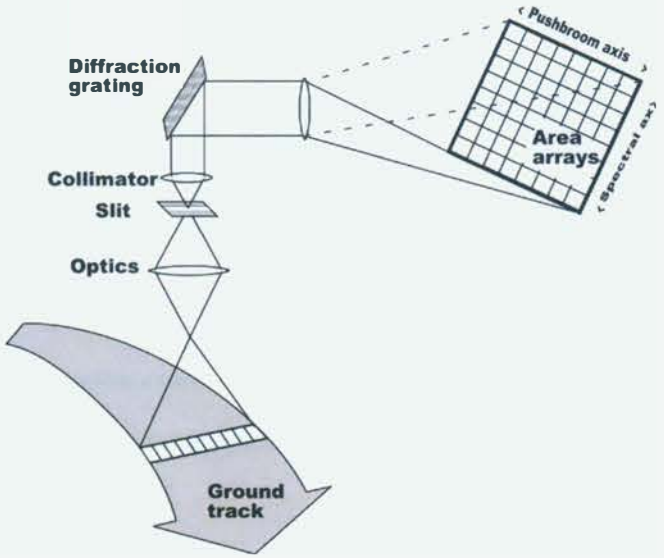


Figure 1.3: Concept of a linear array, or pushbroom sensor.
(2)

travels forward, continuous lines of data are collect to form an image. The Hyperspectral Mapper (HyMap) developed by HyVista corporation is an example of a line scanning Hyperspectral sensor (3). HyMap sensors have been developed which record up to 128 spectral bands covering the visible to the short wave infrared (0.44-2.5 μm) portion of the EM spectrum.

Another type of hyperspectral sensor is a linear array scanner, or pushbroom scanner. These scanners operate much in the same manner as line scanning systems, but they collect data one line at a time, so there is no need for a scan mirror. See Figure 1.3 for a drawing of how a pushbroom scanner collects imagery.

One example of a pushbroom scanner is the Hyperspectral Digital Imagery Collection Experiment (HYDICE) (4). HYDICE is very similar to HyMap in that it operates in the visible to short wave infrared (0.4-2.5 μm), but samples the spectrum with 210 spectral channels. Pushbroom scanners have certain advantages over line scanning systems in terms of their signal to noise ratio (SNR), and they generally have fewer distortions due to the variation in aircraft attitude.

As imaging has progressed, the use of hyperspectral imagery has become advantageous for many applications. Some examples are change detection, vegetation analysis, anomaly detection and target detection. Using hyperspectral imagery, a great amount of data can be collected over a very large area. In the above mentioned example of vegetation analysis, using hyperspectral

imagery to determine crop health can be done much more quickly and cost effectively than the traditional method of sending surveyors out in fields and collecting data.

There are many problems and unresolved issues in analyzing hyperspectral data. In the case of target detection, there are many hyperspectral target detection algorithms. There is no optimal algorithm for every scene. Some algorithms work better than others depending on which materials the scene contains, how cluttered the scene is, etc. If the atmosphere is being modeled using a physics based approach, the sensor must be very precisely radiometrically calibrated. In the case of sensor design, sometimes multiple focal planes are used, or multiple optical fibers on a common focal plane are used to feed separate spectrometers to span a large spectral range. These designs can lead to a spatial misregistration between spectral bands. If the spatial misregistration is significant, the performance of hyperspectral target detection algorithms will suffer.

RIT's Modular Imaging Spectrometer Instrument (MISI) is configured in such a way that two optical fibers at the primary focal plane each feed a separate spectrometer, one in the visible portion of the spectrum (VIS) and the other in the near-infrared (NIR). Since these two fibers cannot occupy the same space, they are located at different locations on the focal plane, and hence there is an inherent spatial misregistration between the VIS and NIR bands.

Appropriate steps may be taken to improve the registration between sets of spectral bands, which should lead to an improvement in target detection performance. The following work will focus on methods to improve spatial registration between bands, as well as investigate the effects various amounts of misregistration have on target detection performance. Since a line scanning system is employed in this research, a scale, rotation and translation of the data will not be an appropriate registration method, as the images are captured over time. More advanced methods using sensor location and orientation parameters recorded during flight will be developed to characterize and compensate for the misregistration. Target detection performance will be evaluated using imagery with this type of misregistration between sets of spectral bands. Different target detection algorithms, both geometric and stochastic, will be considered, and receiving operator characteristic (ROC) curves will be generated to evaluate the performance of such algorithms under varying degrees of spatial misregistration.

Chapter 2

Objectives

The objectives of this thesis are to investigate methods to perform band-to-band image registration on hyperspectral line scanner data, and to determine how band-to-band misregistration in hyperspectral imagery affects different target detection algorithms. An analysis of MISI has been conducted, and a geometric model of the sensor has been made. Several algorithms have been developed that attempt to correct for the misregistration between bands. Recognizing that any resampling method other than nearest neighbor produce spectral mixing between neighboring pixels, we have chosen to maintain the radiometry of the scene by using only a nearest neighbor approach when registering these sets of bands to each other.

Even after two sets of bands are registered to each other, there will still be some amount of error in the registration. This is because the position and orientation of the aircraft, and the digital elevation model (DEM) can never be known exactly. Even if these parameters were known exactly, the points on the ground, as seen by each optical fiber, will never be exactly the same at any 2 pixels, which means a perfect correction is not attainable. This makes target detection difficult especially when dealing with targets that occupy roughly one pixel, or are sub-pixel.

An analysis was conducted to determine the effects misregistration has on various hyperspectral target detection algorithms. Most of the algorithm analysis has been performed using synthetic data. DIRSIG (5)was used to generate a radiometrically accurate scene with 140 vehicles parked against different background surfaces. The DIRSIG imagery was then analyzed with various target detection algorithms to generate receiver operator characteristic (ROC) curves. This would have been difficult to do with real imagery, as there would be very few vehicle target

pixels in the scene to use to generate ROC curves. When attempting to quantify the effects of various amounts of spatial misregistration has on target detection performance, real imagery is not ideal because the amount of misregistration is unknown. Using synthetic data we can easily vary and control the amount of misregistration between sets of spectral bands, and further investigate the effects varying levels of such a misregistration have on different target detection algorithms.

Many interesting questions and concerns arise when dealing with spatial misregistration in hyperspectral imagery. At first glance, it might seem beneficial to simply separate the data into 2 separate sets of bands and apply target detection algorithms to each set of bands separately. This might be a good solution if misregistration is very severe, but what about when misregistration exists at the sub-pixel level? How much misregistration is tolerable for a particular algorithm? Are some types of algorithms more robust than others in the presence of misregistered data? Since spectral mixing occurs because of misregistration, how does the background material surrounding the target affect target detection performance? This thesis will attempt to demonstrate the effects of spatial misregistration up to 0.5 pixel. The research performed will also attempt to show whether it is beneficial to use the full spectrum when data is misregistered, or if it is better to apply target detection algorithms to only one set of bands. Another objective is to determine which types of algorithms are robust in the presence of misregistered data.

Chapter 3

Background

There are many variables which contribute to the final image produced by a sensor. The optical system, the material the detector is made from, the type of scanning mechanism, filters, dispersing elements, amplifier and other electronics are all things that contribute to the fidelity of the final image. Other variables influence the final image which cannot be accounted for in the sensor design. Atmospheric conditions, cloud cover, illumination, and sun angle are all examples of things that affect the captured image, and they all have nothing to do with the sensor. These considerations are all very important especially when dealing with target detection applications using hyperspectral imagery. The material in this background section is meant to give the reader an overview of the details of imaging systems, radiative transfer, data processing, and target detection which must be considered in this research. All information given here is meant to be relevant to the particular tasks set forth to be accomplished in analyzing misregistered data, determining the effects misregistration between spectral bands has on various target detection algorithms, and demonstrating what can be done to improve target detection performance in the presence of such misregistration.

3.1 Remote Sensing Overview

Remote Sensing, as described in the Introduction section, has many practical applications. Just from the introduction and objective section, it can be seen that there are many components to remote sensing which make it very complex. The emphasis here will be on hyperspectral remote sensing, atmospheric compensation, photogrammetry, and target detection algorithms.

Hyperspectral data are collected by sensing at many narrow regions of the electromagnetic (EM) spectrum, or spectral bands. This is not limited to the visible (VIS) region of the spectrum; EM radiation in the any spectral region may be imaged. It is typical for hyperspectral imaging systems to collect in the VIS and infrared (IR) regions of the EM spectrum. The IR region of the spectrum is made up of the near infrared (NIR), short wave infrared (SWIR), mid wave infrared (MWIR) and long wave infrared (LWIR). A HSI system with a great number of narrow spectral bands may be beneficial, especially if targets of interest have unique spectral features in the regions captured by such systems.

Spectroscopy is the field of science concerning the measurement of spectra when matter interacts with or emits electromagnetic radiation. Using spectroscopy, materials can be identified by their unique spectral characteristics and objects can be differentiated from one another. It is easy for humans to visually distinguish many objects based on their spatial structure, but under varying illumination conditions it can be difficult for humans to differentiate objects based on color alone. Humans are also limited to only the VIS region of the spectrum, while there exists useful data in the infrared which can be used to differentiate objects spectrally. This is important for target detection applications. Camouflage designed to look like grass in the visual part of the spectrum may look drastically different from grass in the NIR. In this case, a sensor that collects data in the NIR would be very beneficial when differentiating vegetation from camouflage. It is important for a sensor to be radiometrically calibrated accurately, or to have accurate ground truth in order to convert images to reflectance because spectroscopy relies solely on accurate spectral data.

3.2 Pre-Processing Raw Data

Raw data from a sensor must be processed before it is useful to almost any application. Georeferencing may be performed, which allows points in an image to be related to positions on the ground. Radiance calibration may be performed to relate raw sensor data to radiance values for each band. These techniques do not significantly alter data, but may rearrange, or add to it. When compensating for atmospheric effects, however, the data must be altered. Compensating for atmospheric effects involves predicting reflectance or temperature values from raw sensor data, or observed at-sensor radiance measurements. Further pre-processing may be applied to dimensionally reduce the amount of data. Transformations can be applied to maximize variabil-

ity in an image, or rotate the data such that the resulting transformed image will have ordered in terms of noise. Many bands can be truncated from the transformed images while retaining a large amount of useful data. This results in much smaller file sizes which allows for much easier and faster processing of the data.

3.2.1 Atmospheric Compensation

Sensor data are either in raw digital counts, or in radiance if the sensor has been radiometrically calibrated. The sensor reaching radiance may come from many different sources, due to various reflections, scattering, etc. There is a governing equation for sensor reaching radiance (6) which is given as:

$$L_\lambda = \{E'_{s\lambda} cos(\sigma)\tau_1(\lambda)\frac{r(\lambda)}{\pi} + \epsilon(\lambda)L_{T\lambda} + F\left[E_{ds\lambda} + E_{d\epsilon\lambda}\right]\frac{r_d(\lambda)}{\pi} + (1-F)\left[L_{bs\lambda} + L_{b\epsilon\lambda}\right]r_d(\lambda)\}\tau_2(\lambda) + L_{us\lambda} + L_{u\epsilon\lambda} \quad (3.1)$$

Table 3.1 shows all relevant symbols in this governing equation and what they represent. Most of the discussion in this thesis will involve only reflected electromagnetic radiation. Thus the self emitted terms may be neglected.

Table 3.1: Variables of Sensor Reaching Radiance Equation

$E'_{s\lambda}$	exoatmospheric spectral irradiance $[Wm^{-2}\mu m^{-1}]$
σ	solar zenith angle
$\tau_1(\lambda)$	atmospheric transmission from the sun to the target
$r(\lambda)$	the target reflectance
ϵ	target emissivity
$L_{T\lambda}$	spectral radiance for a blackbody at temperature T
F	fraction of the hemisphere seen by the target
$E_{ds\lambda}$	downwelled irradiance from the sky onto the target
$E_{d\epsilon\lambda}$	downwelled self emitted irradiance from the sky onto the target
r_d	diffuse target reflectance
$L_{bs\lambda}$	background radiance incident on target
$L_{b\epsilon\lambda}$	self emitted background radiance incident on target
τ_2	atmospheric transmission from the sensor to the target
$L_{us\lambda}$	solar upwelled radiance
$L_{u\epsilon\lambda}$	self emitted upwelled radiance

If there is any significant amount of atmosphere between the sensor and the scene (for remote sensing there always is), it is not easy to relate information from the sensor to information in the scene. As can be seen from Table 3.1 there are many variables which contribute to the at sensor

radiance other than the reflectance spectra of objects in the scene. Atmospheric parameters such as temperature, humidity, and aerosols as well as illumination conditions all contribute to the at sensor radiance. Various approaches can be used to compensate for the atmosphere, and convert raw digital counts or radiance values to surface reflectance values. Most attention here will be focused on the empirical line method (ELM) for atmospheric compensation. Other methods are used, but may not be relevant to the work done in this thesis.

3.2.1.1 Empirical Line Method (ELM)

The Empirical Line Method (ELM) uses in-scene data to compensate for atmospheric effects. This method is desirable when working with sensors which have not been radiometrically calibrated because there is no way to extract accurate radiance data from an uncalibrated sensor. Bright and dark calibration panels of known reflectance which fill at least an entire pixel of the image must be present in the scene. These panels should be separated from each other so there is no chance for adjacency effects, or for spectral mixing between image pixels. If no panels are available, bright and dark pixels can be assigned approximate reflectance values (this becomes more difficult when dealing with multi or hyperspectral imagery). ELM assumes equal atmospheric transmittance at every pixel in an image, uniform path radiance, and therefore a spatially, nonvarying linear relationship between radiance (L) and reflectance (r) values. This relationship can be written as

$$L_\lambda = m_\lambda r_\lambda + b_\lambda. \tag{3.2}$$

When given a pixel of known reflectance, and its associated digital count or radiance value, the slope m and offset b can be computed per band. The equation can then be used to convert every radiance value for every pixel in the scene to reflectance. The m term includes the effects of atmospheric transmittance as well instrumental factors. The b term includes dark current offset, as well as atmospheric path radiance. For this method to work in practice, good ground truth is needed, as well as uniform atmospheric conditions.

3.2.1.2 Physics Based Atmospherical Modeling Approach

While setting out calibration panels and measuring reflectances of materials provides a nice approach to converting images to reflectance units, ground truth are often not available. Another

approach to compensating for atmospheric effects is to predict the effects of the atmosphere using physics based modeling. The Low resolution atmospheric transmission model (LOWTRAN) was developed by the U.S. Air Force Research Laboratory to model these effects. LOWTRAN has a relatively poor spectral resolution capability of 20 cm^{-1}. Due to the need for higher spectral resolution, the Moderate resolution atmospheric transmission model (MODTRAN) was developed shortly after, which has a much higher spectral resolution of 2 cm^{-1} (7). MODTRAN, having higher spectral resolution and more accurate band models, quickly became more popular than LOWTRAN, especially for atmospheric modeling for multi and hyperspectral applications. Basically, MODTRAN takes atmospheric parameters such as temperature, humidity, aerosol content, etc. into consideration. LOWTRAN and MODTRAN are both actually radiative transfer models which predict the effects on electromagnetic radiation as it propagates through the atmosphere. Knowing the effects the atmosphere has on electromagnetic radiation, along with sensor altitude, the sensor reaching radiance can be predicted for a reflector on the ground using an atmospheric compensation algorithm (ACA). The fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) algorithm is one MODTRAN-based ACA used in hyperspectral applications when working with regions in the VIS through SWIR portion of the EM spectrum (8).

3.2.2 Dimensionality Reduction

Hyperspectral imaging systems are capable of generating a huge amount of data. This can become problematic as the number of bands, as well as the size of the scene increases. Sometimes when data become too large, some type of data reduction is necessary as a preprocessing step.

3.2.2.1 Principal Component Analysis (PCA)

Principal Components may be used as a dimensionality reduction method. Consider a random vector, $X' = [X_1, X_2, ..., X_p]$, which has a covariance matrix Σ. Linear combinations may be written as:

$$\begin{aligned} Y_1 &= \mathbf{a_1'X} = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \\ Y_2 &= \mathbf{a_2'X} = a_{21}X_1 + a_{22}X_2 + ... + a_{2p}X_p \\ &... \\ Y_p &= \mathbf{a_p'X} = a_{p1}X_1 + a_{p2}X_2 + ... + a_{pp}X_p \end{aligned} \tag{3.3}$$

where a is a weighting coefficient. The variance can be defined as $Var(Y_i) = \mathbf{a_i'\Sigma a_i}$, for $i = 1, 2, ..., p$, where p is the dimensionality of the data. The principal components are the *uncorrelated* linear combinations $Y_1, Y_2, ...Y_p$ which maximize the variances. (9) The main assumption is that variance is proportional to information in the scene. By transforming a scene to a space where the variance is ordered, bands with low variance can be truncated and this space may be used to reduce the size of the data. Small variances may also be zeroed out, then the data can be back transformed to reduce noise.

3.2.2.2 Minimum/Maximum Noise Fraction (MNF)

The Minimum Noise Fraction (MNF) transform attempts to order data in terms of signal to noise ratio. By employing this method, noise sources due to thermal variation, photon noise, detector noise, etc. can be reduced. Consider an image $X = S + N$, where S represents signal and N represents noise. The covariance representation for such an image may be expressed as

$$\Sigma_X = \Sigma_S + \Sigma_N \tag{3.4}$$

Using similar methods used for PCA, the data may be transformed and ordered in terms of SNR (10). In the literature, the transformation to the new space is called Minimum Noise Fraction, or Maximum Noise Fraction, depending on how the data are ordered.

3.3 Target Detection Algorithms

3.3.1 Spectral Angle Mapper (SAM)

One of the simplest hyperspectral target detection algorithms is the spectral angle mapper (SAM) developed by Boardman and Kruse (11). SAM computes the angle between two vectors.

If two vectors are very similar, the angle between them will be small. In vector form, SAM can be written as

$$T_{SAM} = \frac{d^T x}{(d^T d)^{1/2}(x^T x)^{1/2}} \tag{3.5}$$

where d is the target vector and x is the vector of the pixel of interest. Notice that equation 3.5 is normalized by the magnitude of both the target and pixel of interest vectors. This ensures all values for the algorithm will fall between 0 (orthogonal to the target vector, least target-like) and 1(a perfect target match).

SAM can be used with any type of data but typically assumes a hyperspectral image which has been converted to units of reflectance. This is because target signatures are usually known in reflectance units. SAM is a good baseline detection algorithm as it is very computationally inexpensive. It depends only on spectral shape and is invariant to the magnitude of the vectors. This could become a problem when attempting to differentiate between materials which have similar spectral shape, but are dissimilar in magnitude. More elaborate algorithms account for both spectral shape and magnitude, as well as scene statistics to increase detection rates, and decrease the number of false alarms.

3.3.2 Geometrical (Structured) Background Detectors

Scenes contain background materials which must be characterized when applying geometric target detection algorithms. Different background materials are typically referred to as end-members. They can be extracted from imagery using a clustering technique to classify the scene background. Several clustering techniques exist, and one well known algorithm called K-means is described here. K-means is an unsupervised, iterative procedure used in remote sensing applications to classify pixels based on their spectrum. It is assumed that a hyperspectral image can be decomposed into K classes, where K is a known positive integer value. The first step involves artbitrarily grouping pixels into K initial clusters, and computing the center of each cluster. In the second step, each pixel in the image is assigned to a cluster based on the closest Euclidean distance to a cluster mean. In the third step the cluster centers are recomputed. Steps (2) and (3) are repeated until the number of pixels changing classes becomes sufficiently small, or the maximum number of iterations set by the user has been reached.

Other algorithms exist, such as the Gaussian Maximum Likelihood (GML) classifier which separates classes based on their statistical, rather than Euclidean distance. For all work done

in this thesis, K-means was used as the clustering algorithm for geometric target detection algorithms for the sake of simplicity and consistency.

Geometric and Stochastic target detection algorithms rely on a binary hypothesis test to determine whether a target occupies a pixel in a scene. For a geometric algorithm, this test is written as

$$H_0 : x = Ma_b \tag{3.6}$$

$$H_1 : x = da_t + Ma_b \tag{3.7}$$

where H_0 is the null hypothesis when there is no target present, and H_1 is the hypothesis when some amount of target is present. M represents a matrix of background endmembers, and d, again represents the target. Abundances are denoted by the vectors a_b (background abundance) and the scalar a_t (target abundance). The matrix of background endmembers, M, is used in such a way to suppress the background, usually by projecting vectors in such a way that the background endmembers are nullified.

3.3.2.1 Orthogonal Subspace Projection (OSP)

The Orthogonal Subspace Projection algorithm (OSP) developed by Harsani and Chang is one type of matched filter based on modeling the background geometrically (12). This approach basically projects each pixel vector onto a subspace orthogonal to the undesired signatures (background signatures). To do this, an operator, P, is used to make the projections of pixel vectors onto the subspace mentioned above. P can be defined as

$$P = (I - MM^{\#}) \tag{3.8}$$

Where I is the identity matrix, M is a matrix of background endmembers and $M^{\#}$ is called the pseudo inverse of M and can be written as

$$M^{\#} = (M^T M)^{-1} M^T. \tag{3.9}$$

The operator, P can be applied to every pixel vector in an image. P operating on a pixel

where no target is present, according to the test hypothesis in equation 3.6, results in

$$\begin{aligned} P_x &= PMa_b \\ &= (I - MM^{\#})Ma_b \\ &= (M - MM^{\#}M)a_b \\ &= (M - MI)a_b \\ &= 0 \end{aligned}$$

whereas P operating on a pixel with some abundance of target, according to the test hypothesis in equation 3.7 results in

$$\begin{aligned} P_x &= Pd_{a_t} + PMa_b \\ &= Pd_{a_t}. \end{aligned}$$

It is assumed that target vectors are not background-like and so resulting values using this method are high for target-like pixels and low for background-like pixels. The results from this algorithm are often normalized and can be written as

$$T_{OSP}(x) = \frac{d^T Px}{d^T Pd} \tag{3.10}$$

3.3.2.2 Adaptive Subspace Detector (ASD)

Another geometric target detection algorithm is the Adaptive Subspace Detector (ASD) (13). This algorithm is usually used when a target subspace is desired, as opposed to using a single target vector. It is similar to OSP, but is unique due to the introduction of this target space. The algorithm is defined below

$$T_{ASD}(x) = \frac{x^T(P_b - P_S)x}{x^T P_S x} \tag{3.11}$$

where P_S is the matrix of target vectors concatenated with the matrix of background endmembers. P_S may be formed by taking multiple measurements of the same target, or by modeling what the target may look like under various circumstances. More on this topic is discussed in

section 3.4 regarding the invariant algorithm and physics based modeling of the atmosphere.

3.3.3 Unstructured (statistical) Target Detection Algorithms

It is sometimes useful in the case of an unstructured background to use statistical target detection algorithms. These methods rely on background characterization through statistics rather than background endmember selection. The binary hypothesis test for statistical target detection algorithms is of the form:

$$H_0 : x = \nu \tag{3.12}$$

$$H_1 : x = da_t + \nu \tag{3.13}$$

here, again, H_0 is the null hypothesis, H_1 is the hypothesis of some target present, x is the spectrum of a test pixel, d is a the target spectrum, a_t is a fractional abundance and ν is a background spectrum.

3.3.3.1 Generalized Likelihood Ratio Test (GLRT)

Johnson (9) provides a solution to the generalized likelihood ratio test (GLRT) for the binary hypothesis tests 3.12, 3.13.

$$T_{GLRT}(x) = \frac{(d^T \Sigma_b^{-1} x)^2}{(d^T \Sigma_b^{-1} d)(1 + x^T \Sigma_b^{-1} x)} \tag{3.14}$$

where Σ_b is the background covariance.

If some assumptions are made about the GLRT, the Adaptive Coherence Estimator (14) can be derived (15). The background covariance, in practice, is usually not known. It can be estimated, however, using a sample of pixels. If a small sample of pixels is chosen to represent the background, the estimated background covariance becomes

$$\hat{\Sigma}_b = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T \tag{3.15}$$

where N is the number of pixels in the sample data. The GLRT can now be written as

$$\hat{T}_{GLRT}(x) = \frac{(d^T \hat{\Sigma}_b^{-1} x)^2}{(d^T \hat{\Sigma}_b^{-1} d)(1 + \frac{1}{N} x^T \hat{\Sigma}_b^{-1} x)}. \tag{3.16}$$

As long as the number of training samples, N, remains relatively small, the $(1 + \frac{1}{N} x^T \hat{\Sigma}_b^{-1} x)$ term reduces to $(x^T \hat{\Sigma}_b^{-1} x)$ (15). This brings the approximation of the GLRT to

$$\hat{T}_{ACE}(x) = \frac{(d^T \hat{\Sigma}_b^{-1} x)^2}{(d^T \hat{\Sigma}_b^{-1} d)(x^T \hat{\Sigma}_b^{-1} x)} \tag{3.17}$$

which is the Adaptive Coherence Estimator. ACE is invariant to scale changes in the test and training data, and can be shown to be the cosine of the angle between the test pixel and the target subspace when a whitened subspace is used.

Another variation of the GLRT is constrained energy minimization (CEM) (16). The CEM works by producing an estimate of the fractional abundance of a target material for each pixel in a scene. CEM can be written as

$$\hat{T}_{CEM}(x) = \frac{(d^T \hat{\Sigma}_b^{-1} x)}{(d^T \hat{\Sigma}_b^{-1} d)}. \tag{3.18}$$

3.4 The Forward Model

Great effort has been put forth to radiometrically calibrate hyperspectral sensors. Radiometric calibration is essential when ground truth is not available. When radiative transfer codes are being used to determine sensor-reaching radiance values, it is often called forward modeling. Many obstacles must be overcome when using forward modeling. As mentioned above, sensors need to be very accurately radiometrically calibrated; error or large uncertainty in the calibration can cause large discrepancies when comparing sensor radiance values to those predicted using radiative transfer code. Another obstacle is the radiative transfer code itself. Atmospheric physics based models are non-linear and can become extremely complex. If all these obstacles can be overcome, forward modeling has some very attractive advantages over atmospheric compensation approaches, especially when dealing with target detection.

The basic idea of atmospheric compensation is to get rid of atmospheric effects in the image. This can be done using an ELM, but has its limitations. If the atmospheric or illumination conditions are not uniform across the image, targets with similar reflectances will appear different to the sensor at different places in the scene. Also, the calibration panels used for an ELM are most often out in the open where they reflect direct sunlight as well as downwelled irradiance. Even when illumination conditions are uniform, targets in shaded regions will not

be well represented by atmospherically compensated data, which makes target detection very difficult for objects in this case. Techniques can be included in the forward modeling approach which incorporate variability in atmospheric conditions as well as target under various illumination conditions. Work has also been done to predict target signatures of objects at different orientations in a scene, such as targets on the side of a hill.

3.4.1 The Invariant Method

In a 1999 experiment Glenn Healey and David Slater proposed a target detection method using forward modeling which would be invariant to illumination conditions (17). In their method they assumed the dominant sources of sensing reaching radiance were from (1) reflected solar illumination, (2) reflected sky illumination, and (3) upwelled radiance. Adjacency illumination from background objects was considered to be negligible. The HYDICE sensor, which covers the spectral region from 0.4-2.5μm, was used to collect imagery so thermal radiance was also not considered. Four atmospheric profiles were used: U.S. standard, midlatitude summer, midlatitude winter, and tropical. Four watervapor profiles were used, as well as four different profiles for the atmospheric gases O_3, CH_4, N_2O and CO. Four aerosol profiles were also used: rural, urban, maritime and desert. Eight solar zenith angles were also considered, varying from 5 to 75 degrees. In addition to all these parameters, the target would be considered under 2 illumination conditions: (1) both solar and sky illumination, and (2) sky illumination only. Using these parameters, a combination of 28,672 conditions are possible, 17,920 of which are physically realizable. This means 17,920 spectral curves may be computed to characterize the target of interest under these various conditions.

Because of the high dimensionality of hyperspectral datasets, dimensionality reduction techniques are often employed, as has been discussed in Section 3.2.2. In this particular case, 210 spectral bands were recorded. A smaller number of basis vectors was desired which would satisfactorily represent the background and target subspaces. To demonstrate this symbolically, let W represent the number of spectral bands and C represent the total number of conditions we would like to represent. A material under C different conditions can be represented in radiance space as $L_1(\lambda), L_2(\lambda), ..., L_C(\lambda)$. Any radiance curve, L_i can be written as

$$L_i = (L_i(\lambda_1), L_i(\lambda_2), ..., L_i(\lambda_W))^T \tag{3.19}$$

for $1 \leq i \leq C$. L_i can be approximated using fixed orthonormal basis vectors $m_j = (m_j(\lambda_1), m_j(\lambda_2), ..., m_j(\lambda$ for $1 \leq j \leq N$ and weighting coefficients a_{ij} which depend on the particular set of conditions when L_i was recorded. This approximation can be written as:

$$L_i \approx \sum_{j=1}^{N} a_{ij} m_j \tag{3.20}$$

The accuracy of this approximation may be found by taking the squared error

$$E_i(N) = \left(L_i - \sum_{j=1}^{N} a_{ij} m_j \right)^2 \tag{3.21}$$

Under the total set of C conditions, the total squared error using N basis vectors is expressed as

$$E_T(N) = \sum_{i=1}^{C} E_i(N) \tag{3.22}$$

Techniques based on the singular value decomposition can be used to find the orthonormal basis for any N which minimizes the total squared error $E_T(N)$. This was done for 498 materials in the USGS database to determine how many basis vectors can accurately approximate a set of spectral vectors. The maximum normalized error when using 9 basis vectors was 0.0002. This suggests many materials may be accurately represented under various conditions with far fewer basis vectors than the number of spectral bands used to collect the imagery.

3.4.1.1 Target Detection Experiment

Healey and Slater used a scene which had an open field with several targets of known reflectance placed both in the middle of an open field, some targets in the shade by a treeline, as well as some targets placed further into the trees where they could be seen by the sensor, but were illuminated only by downwelled radiance. It was first demonstrated that simple algorithms such as SAM would not perform for targets under different illumination conditions. SAM depends only on the spectral angle, so the magnitude of the spectral vectors does not matter. An in scene spectrum was collected for a target in the middle of the field, as well as for a target in a shaded region. The spectra were peak-normalized and plotted on a common set of axes. The spectra can be seen in Figure 3.1. It can clearly be seen that under different illumination conditions the targets appeared very spectrally different. This is because one of targets is lacking the direct

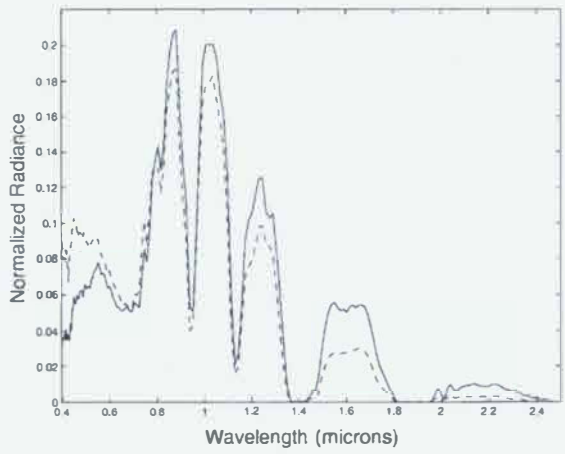


Figure 3.1: Spectra of targets in direct sunlight and in shadow. The dashed line represents the shaded target and the solid line represents the target directly illuminated by the sun.

term, and is illuminated only by downwelled radiance. Notice the target lacking the direct term has a higher normalized radiance in the blue portion of the electromagnetic spectrum. SAM was applied to the image using in-scene spectra from a non-shaded target. While other non-shaded targets all got high returns, the shaded targets were not at all easily detected.

The invariant approach was tried next, using all the radiance vectors under various conditions described above. An algorithm was developed using maximum-likelihood parameter estimates. Normalized radiance vectors were described as

$$\hat{L} = \sum_{j=1}^{N} \alpha_j m_j + \eta \tag{3.23}$$

This equation is similar to equation 3.20, but here η is a residual which is modeled as a zero-mean Gaussian random vector with small covariance elements since the approximate error is small. The likelihood that a radiance vector $\hat{L}$ represents a material of reflectance $R(\lambda)$ under the parameter values $\alpha_1, \alpha_2, ..., \alpha_N$ is computed as

$$P(\hat{L}|R, \alpha_1, \alpha_2, ..., \alpha_N) = \frac{1}{(2\pi)^{0.5W}|\Sigma|^{0.5}} \exp -0.5 D^T \Sigma^{-1} D \tag{3.24}$$

where N is the number of basis vectors and Σ represents the covariance matrix of η here, and

$$D = \hat{L} - \sum_{j=1}^{N} \alpha_j m_j \tag{3.25}$$

A likelihood was applied for every pixel in the scene and the image was thresholded to identify targets. It was shown this method approached the performance of SAM for directly illuminated targets at higher false alarm rates. The interesting thing about this approach is targets in shadow were correctly identified at low false alarm probabilities, while the probability of detection using SAM never went above 60% at any false alarm rate. These results show that when the particular conditions a target is under are not known, incorporating the variability of the target signature into the model can improve target detection algorithm performance.

3.4.1.2 Implications of Invariant Method

The invariant method is important because it performs well dealing with uncertainties and unknowns. Various illumination levels and atmospheric conditions can be used to create a target space using multiple target signatures based on the conditions. It might be possible to apply this same principal in dealing with unknowns in the amount of spatial misregistration in a hyperspectral sensor working in either radiance space or reflectance space. If the target signature is known, and the background material surrounding the target is known, multiple target signatures can be constructed using a linear mixture model in only 1 set of spectral bands. The linear mixture would depend on the amount of spatial misregistration.

3.5 Image Registration and Photogrammetry

Once images have been acquired, they are often georectified for mapping or GIS purposes. Registration is also important if data sets are to be fused together. For example, a hyperspectral image cube may be aquired, and LIDAR data may also be available for the same scene. By combining the two datasets, more data is readily available which can be extracted from the images. Features such as height information that would not otherwise be available from just the hyperspectral data would now be present in the combined data set. Combining the two data sets requires the images to be precisely registered to one another.

Georeferencing involves applying a transformation to an image to align ground points with ground coordinates. Several techniques can be used. A two dimensional conformal coordinate transformation consists of only a scale change, rotation and translation of data. The true shape of the collected data is preserved in this case. Control points in the image are used to solve for these parameters. If more than two control points are used, the parameters may be solved for

using a least squares solution. Given the ground locations, A, B of two points, a, b for a two dimensional image, we can solve for the scale using equation 3.26

$$s = \frac{AB}{ab} = \frac{\sqrt{(E_B - E_A)^2 + (N_B - N_A)^2}}{\sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}} \tag{3.26}$$

where A, B have been broken up into their east (E) and North (N) components, and a, b have been broken up into their x and y components. A two dimensional coordinate transformation is used to determine the angle of rotation, θ. A description of how a point, C is rotated in the image is given in equation 3.27

$$\begin{aligned} E'_C &= X'_C \sin(\theta) + Y'_C \cos(\theta) \\ N'_C &= X'_C \sin(\theta) + Y'_C \cos(\theta) \end{aligned} \tag{3.27}$$

where X'_C, Y'_C are the axes of the coordinate system rotated relative to E'_C, N'.

A translation may be accomplished by shifting the data a certain amount in the N and E directions. That is,

$$\begin{aligned} E_C &= E'_C + T_E \\ N_C &= N'_C + T_N \end{aligned} \tag{3.28}$$

here, T_E and T_N are the translation coefficients. They are calculated as

$$\begin{aligned} T_E &= E_A - E'_A \\ T_N &= N_A - N'_A \end{aligned} \tag{3.29}$$

where a hypothetical control point A is used here.

Another common type of coordinate transformation is the 2 dimensional affine coordinate transformation. This transform is very similar to a conformal transform, with the exception of two important differences. The Affine transformation allows for independent scale factors in the x and y directions, as well as non-orthogonal axes. In this case, scale factors are s_x and s_y such that

$$\begin{aligned} x' &= s_x x \\ y' &= s_y y \end{aligned} \tag{3.30}$$

Several geometries are possible to cause non-orthogonalities in images. None will be mentioned here, but ξ will denote a non-orthogonality operator to adjust for the effect in an image.

$$\begin{aligned} x' &= \xi_x(x') \\ y' &= \xi_y(y') \end{aligned} \tag{3.31}$$

The rotation step for this method is the same as in equation 3.27, and the translation step is the same as in equation 3.28. Other methods such as higher order polynomials may be used for image registration, but are not necessary here.

In chapter 4 collinearity condition equations will be developed which consider the 6 external parameters (roll, pitch, yaw of the aircraft, as well as its location in 3-D space). The collinearity equations, also called projective equations, can use knowledge of all 6 of these parameters to rectify an image. Although these equations are normally used when dealing with framing arrays, we will find them useful in constructing a geometric model of a line scanner. After the scanner geometric model is created, these equations will also be useful in projecting pixels onto the earth to compute the ground location as seen by each pixel in an image.

When transforming two images to one another, the data in at least one image must be manipulated. The type of transformation tells how points in one image should be changed to match the points in another image. Digital images are made up of discrete pixels. These digital images need to be rotated, translated and scaled to achieve registration. Since quantized blocky images can't be rotated or translated by fractional pixel amount, some sort of interpolation must be employed to manipulate them. A nearest neighbor interpolation is the fastest and simplest interpolation method. Consider an image consisting of pixels on a rectangular grid. For a pixel superimposed a fractional distance between pixels, the distance to each neighboring pixel is computed. The value assigned to the superimposed pixel is the value of the closest neighboring pixel. Interpolating using nearest neighbor may produce jagged or blocky looking images. A

bilinear interpolation is also a popular interpolating choice which often produces smoother looking images than nearest neighbor. Again, for a pixel superimposed a fractional distance between pixels, the distance to each neighboring pixel is computed. A weighted average is computed based on the values and distances to all 4 neighboring pixels. The value assigned to the superimposed pixel is this resulting weighted average. A bicubic interpolation, or cubic convolution is another popular resampling technique. Based on sampling theory, if an image is sampled above Nyquist, it can be almost perfectly reconstructed using ideal sinc interpolation. Using the ideal sinc interpolator for digital image reconstruction would require using the sinc function, which has infinite support, and an image with an infinite number of rows and columns. This would, in turn require infinite running time to compute. Since this is not a practical solution, an image with a finite number of rows and columns may be faithfully reproduced using a cubic spline which nearly approximates the sinc function. All 3 of these methods are commonly used when interpolating images on a rectangular grid.

The techniques mentioned above work well for framing array data, since each pixel is collected simultaneously. When lines of data, or individual pixels are collected over time, these methods break down, as the data is collected from different locations, and with the sensor at various orientations. Akima writes on interpolation methods using *local* rather than *global* procedures for irregularly distributed points on an $x - y$ plane (18). Since line scanner imagery is not collected simultaneously, but over time, local procedures work well to perform transformations. An elastic registration method is required in this case. The method proposed is that a plane is divided into a number of triangular cells, with the requirement that each vertex of a triangle is a data point. Triangulation is chosen by a method suggested by Lawson (19). Lawson's method basically finds quadrilaterals in the $x - y$ plane which have each internal angles less than π. These quadrilaterals can be divided up into 2 triangles. The triangles should be chosen such that the partitioning maximizes the minimum interior angle. A fifth-degree polynomial is used to interpolate data points inside the triangles.

$$z(x,y) = \sum_{j=0}^{5} \sum_{k=5-j}^{5} \P_{jk} x^j y^k \tag{3.32}$$

There are 21 coefficients, $\P_{jk}$ which are determined by the following constraints

• The bivariate function $z(x,y)$ and its first and second partial derivatives in x and y agree at all triangle vertices.

• The partial derivative of the function differentiated in the direction perpendicular to each side of the triangle is a polynomial of degree three, at most, in the variable measured in the direction of the side of the triangle.

Wiemker writes that because this interpolating method is continuous and smooth, it is good for image registration when using line scanner data (20). This method is also computationally very fast compared to using global elastic methods.

This method can be used to directly resample images in the case of georectification, however when dealing with multi or hyperspectral imagery, only nearest neighbor sampling should be used. This is because methods other than nearest neighbor produce spectral mixing between neighboring pixels. This approach is ideal for georeferencing line scanner data, as a smooth interpolation is achieved, but the radiometric accuracy is disturbed in the process. Akima's method, however, is an indirect sampling method which gives the location of pixels in an image. This approach would be very useful for finding ground locations for both sets of bands in hyperspectral imagery, and then using a nearest neighbor method to perform the actual registration.

An alternative method for image registration from line scanner data is proposed by Wanpeng Zhang (21). This method utilizes onboard flight parameters. By knowing the aircraft position and attitude for each collected pixel, the position on the ground for the given pixel can be found. Given the ground locations for each pixel, again a nearest neighbor interpolation may be performed. This approach will be discussed in detail in Chapter 4.

3.6 Modular Imaging Spectrometer Instrument (MISI)

The Modular Imaging Spectrometer Instrument (MISI) was designed, built and maintained by staff and students at the Rochester Institute of Technology (RIT). MISI is a hyperspectral line scanner equipped with a fold mirror, a 6 inch rotating scan mirror and a Cassegrain telescope which has a focal length of 0.5m. Separate focal planes are present in MISI's design for VNIR, SWIR, MWIR, and LWIR detectors, but most of our attention will be focused on the primary focal plane where the VIS and NIR spectra are collected. At the time this thesis was written, the instantaneous field of view (IF●V) for the VIS and NIR channels was calculated to be roughly 2.54 milliradians based on the focal length of MISI's optical system and the diameter of optical fibers at the primary focal plane. This will be discussed in greater detail in Chapter

4. Table 3.2 provides specifications of the instrument. Based on the MISI's IFOV, the GIFOV value for the VIS and NIR channels at 2000′ is 5.08′ but has been rounded to 5′ in Table 3.2.

Table 3.2: MISI specifications

Spectral Band	Center Wavelength (μm)	# Channels	$\Delta\lambda$	GIFOV @ 2000' AGL (feet)
VIS	0.41 - 0.75μm	35	0.012μm	5'
NIR	0.74 - 1.02μm	35	0.010μm	5'
SWIR	1.26μm	1	0.11μm	4'
SWIR	1.65μm	1	0.38μm	4'
SWIR	2.03μm	1	0.65μm	4'
MWIR	3.65μm	1	0.9μm	4'
LWIR	9μm	1	2μm	4'
LWIR	11μm	1	2μm	4'
LWIR	11.5μm	1	2.1μm	4'
LWIR	11μm	1	6μm	2'

MISI's line scanner rotates to collect incident light and reflect it onto three separate focal planes in the across-track direction. The field of view (FOV) of the instrument is ±45°, which gives a 4km FOV when flying at 2000m above ground level. The long wave detectors onboard the instrument can be calibrated between scan lines by having the detectors view on-board visible and thermal calibration sources (22). A diagram of the instrument can be seen in Figure 3.2.

MISI is also equipped with some important equipment used for measuring the instrument's location and orientation during flight. A Crossbow VG400 series digital measurement unit (DMU) is a motion and attitude sensing unit which provides useful information about the instrument orientation due to the dynamics of the aircraft. Roll, pitch and yaw data are recorded by this instrument, although the instrument was set up to record absolute heading instead of yaw on some previous collected data sets. A GPS unit onboard also records the location of the instrument for every line of imagery. Although GPS locations are recorded at every line, the unit itself updates the measurements at longer intervals. Observations of the recorded GPS data show the unit makes measurements every 20-30 lines. Due to the long time gaps in the recorded GPS data, an interpolation method must be employed to estimate the position of the aircraft at every line to project each pixel onto the earth and compute a ground location. Since the position of the aircraft is not known at every line, the accuracy of pixel ground locations is compromised. Ground locations may be computed more accurately if a higher quality GPS unit is installed.

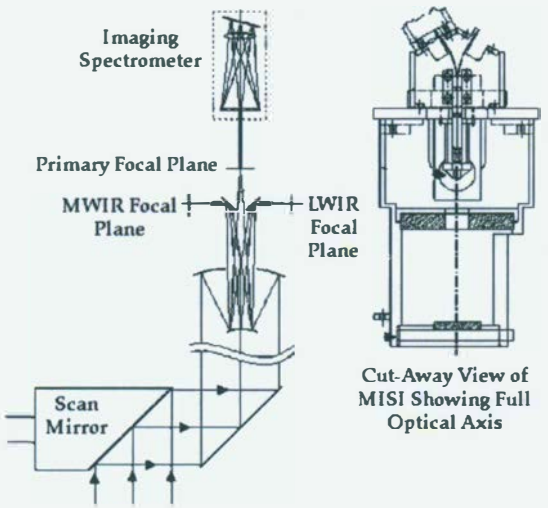


Figure 3.2: Diagram of the MISI instrument and its components.

Chapter 4

MISI Misregistration

MISI has an inherent registration issue between the VIS and NIR bands. Each of these spectrometers is fed by their own optical fiber at the primary focal plane. Since these two fibers cannot occupy the same space, they are placed at different locations on the primary focal plane. The fact that these two fibers are not in the same exact location is what causes the misregistration.

From geometrical optics, we know that rays passing through the optical center of a lens may be drawn as straight lines (23). Applying this principle, we can trace rays through MISI's optical system to determine more information about the registration problem. The fibers are located on opposite sides of the center channel as shown in Figure 4.1. The separation between the fibers has been measured in the lab to be roughly $2.54mm$, and the fibers themselves have been measured to be roughly 1.27mm in diameter (24). Using the separation between fibers on the focal plane we can determine the angular difference between these two rays as they pass through the center of a lens. Looking at Figure 4.2, we can easily notice that the angle can be written as equation 4.1

$$\theta_{VIS/NIRseparation} = \sin^{-1}\left(\frac{s}{f}\right) \tag{4.1}$$

Figure 4.1: Illustrating the locations of the optical fibers on MISI's focal plane.

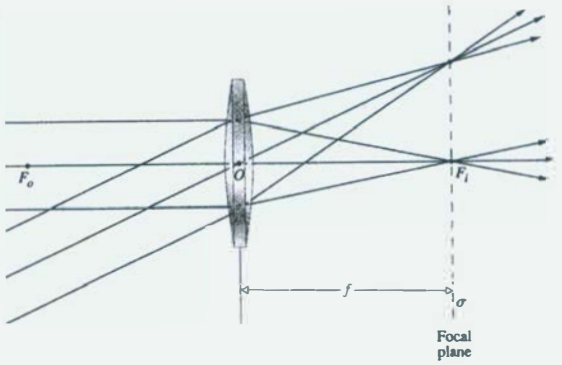


Figure 4.2: A diagram of the MISI instrument and its components.

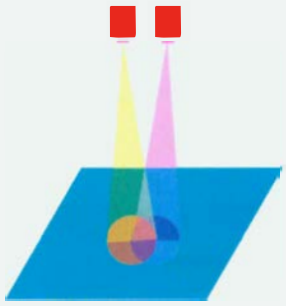


Figure 4.3: Illustration of the effects of misregistration on the ground due to optical misalignment.

where s is the distance of separation and f is the focal length of the lens. Using values of $s = 2.54mm$ and $f = 0.5m$, we get an angle of separation of 5.08 milliradians. This translates to a minimum distance of separation on the ground of $5.08m$ at nadir, at a flying height of 1000m AGL. This effect is illustrated in Figure 4.3. The IFOV (in radians) of the VIS and NIR optical fibers can also be computed using equation 4.1. Using a fiber diameter of $1.27mm$ the IFOV is roughly 2.54 milliradians, or 2.54 meters at a flying height of 1000m. To avoid confusion, this IFOV should only be used to compute the ground spot size for a given pixel. The amount of registration in terms of numbers of pixels must be computed using equation 4.1, MISI's FOV, and the number of pixels recorded per line. Details on this will be discussed in Chapter 5.

When looking at a MISI image, we notice features in the NIR channels appear to be positioned higher and to the right relative to the same features in the VIS channels when observing the left side of the image. We also notice features in the NIR channels appear to move lower and to the right relative to the same features in the VIS channels when observing the right side of

the image. This fact leads one to believe the angle of separation between the two optical fibers is a function of the scan mirror angle. This means there is no simple correction for the problem without a more thorough knowledge of the instrument. The first step to understanding and correcting this problem is to create a geometric model of the sensor. Once a geometric model is created, we can determine the location the sensor is pointing to on the ground for each pixel in an image for each set of bands.

A geometric model of MISI was created with the help of colleagues (25). A ray tracing technique was used to trace rays from the center, VIS and NIR optical fibers through MISI's optical system to determine the direction of the rays leaving the rotating scan mirror. This allows us to tell where each spectrometer is pointing relative to each another. For consistency, a right handed coordinate system has been chosen such that the x direction was the direction of flight of the aircraft, y pointed in the direction of the left wing of the aircraft, and z pointed upward. All vectors propagated throughout the optical system use this coordinate system. A general solution has been developed as follows:

We begin with the focal plane lying on the XY plane of our coordinate system with the center channel located at X, Y=0. Let the optical system consist of a primary focal plane, a telescope of focal length f, a stationary fold mirror, and a rotating scan mirror. The normal vector to the stationary fold mirror can be written as

$$\hat{N_1} = \left[-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right], \tag{4.2}$$

Additionally, the scan mirror rotates from -45 deg to +45 deg. This means a line of data is scanned along the Y axis from the negative Y direction toward the positive Y direction. A normal vector to the scan mirror is described as

$$\hat{N_2} = \left[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\cos\theta, -\frac{\sqrt{2}}{2}\sin\theta\right] \tag{4.3}$$

where θ is the rotation angle of the scan mirror as mentioned above. Drawing a ray from the offset channel on the focal plane through the optical center of MISI's telescope gives us:

$$\hat{A} = \frac{[-x_o, -y_o, -f]}{\sqrt{x_o^2 + y_o^2 + f^2}} \tag{4.4}$$

where x_o, y_o denotes the position of the offset fiber on the focal plane. This ray can be traced

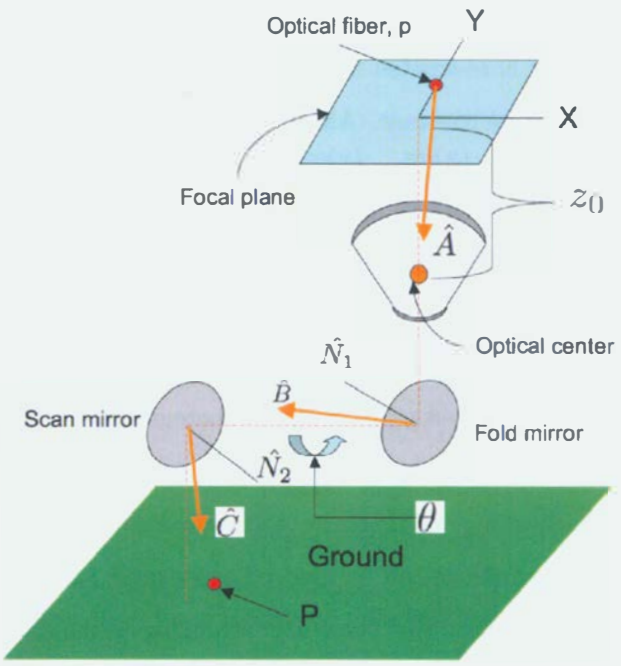


Figure 4.4: Ray tracing through MISI's optical system.

out until the first fold mirror is reached. $\hat{A}$ can be reflected off the fold mirror by rotating it about the fold mirror's normal vector. The resulting reflection is given as:

$$\hat{B} = 2\left[-\hat{A}\cdot\hat{N}_1\right]\hat{N}_1 + \hat{A} \tag{4.5}$$

Tracing the ray out until the scan mirror is reached, we perform the same type of reflection about the scan mirror. Remembering N_2 is a function of θ, we can write the resulting reflection off the scan mirror as:

$$\hat{C} = 2\left[-B\cdot\hat{N}_2\right]\hat{N}_2 + \hat{B} \tag{4.6}$$

A diagram of the ray tracing described in equations 4.4 through 4.6 can be seen in Figure 4.4. These equations tell us the direction of incoming radiation reaching an offset fiber. When the sensor is looking at the ground from above, it is necessary to know the ground location that the fiber is looking at. To determine this, it is necessary to develop a world coordinate system in which we can place and orient the sensor. We can let X, Y, Z represent a set of axes in this world, ground-based coordinate system. T_x, T_y, and H can represent the X, Y, and Z coordinates of the sensor respectively. A vector C_x, C_y, C_z will be used to represent the $\hat{C}$ vector

determined above. This vector is re-written here because a transformation will be necessary, as the aircraft attitude will not be constant. Aircraft roll, pitch, and yaw must be taken into account before the vector can be projected onto the earth to obtain the ground location at each sample.

Colinearity equations have been used traditionally in photogrammetry to correct for aircraft orientation when dealing with data obtained from framing cameras. A rotation matrix can be developed to determine the locations of points in a transformed coordinate system. In the case of MISI, these colineary equations can be used on a per-pixel basis to correct for sensor orientation, and determine the position that the sensor is looking on the ground for every pixel. This can be done for each pixel in a scene for each spectrometer onboard. Pixels from one set of bands can then be paired with pixels from the other set of bands based on their ground locations.

As stated above, the colinearity equations take roll, pitch, and yaw into account. In terms of aircraft orientation, these parameters each define a rotation. The order the operators are applied does matter and the order is: roll (ω), pitch (ϕ), then yaw (κ). The order here is important, as mixing the order of the operations will produce a different result. Orienting the aircraft with the nose pointing in the $+x$ direction, and the left wing in the $+y$ direction, the roll operator may be written as:

$$roll = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & \sin(\omega) \\ 0 & -\sin(\omega) & \cos(\omega) \end{bmatrix} \tag{4.7}$$

the pitch operator may be written as:

$$pitch = \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \tag{4.8}$$

and the yaw operator may be written as:

$$yaw = \begin{bmatrix} \cos(\kappa) & \sin(\kappa) & 0 \\ -\sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.9}$$

By multiplying these operators together, a rotation matrix, M is derived. That is,

$$M = \begin{bmatrix} m11 & m12 & m13 \\ m21 & m22 & m23 \\ m31 & m32 & m33 \end{bmatrix} \tag{4.10}$$

where m's listed below have been substituted for the resulting coefficients from the matrix multiplication (26).

$$\begin{aligned} m_{11} &= \cos(\phi)\cos(\kappa) \\ m_{12} &= \sin(\omega)\sin(\phi)\cos(\kappa) + \cos(\omega)\sin(\kappa) \\ m_{13} &= -\cos(\omega)\sin(\phi)\cos(\kappa) + \sin(\omega)\sin(\kappa) \\ m_{21} &= -\cos(\phi)\sin(\kappa) \\ m_{22} &= -\sin(\omega)\sin(\phi)\sin(\kappa) + \cos(\omega)\cos(\kappa) \\ m_{23} &= \cos(\omega)\sin(\phi)\sin(\kappa) + \sin(\omega)\cos(\kappa) \\ m_{31} &= \sin(\phi) \\ m_{32} &= -\sin(\omega)\cos(\phi) \\ m_{33} &= \cos(\omega)\cos(\phi) \end{aligned}$$

A forward transformation will produce points, or in our case vectors, in the transformed space, that is, in the sensor's frame of reference. We actually want the opposite, to know where the sensor is pointing with respect to the world coordinate system so each pixel can be projected to the ground. The rotation matrix is an orthonormal matrix, which has the property that its inverse is equal to its transpose. This means

$$M^{-1} = M^T \tag{4.11}$$

We can now use M^T to operate on $\hat{C}$.

$$\hat{C}' = C'_x, C'_y, C'_z = M^T \hat{C} \tag{4.12}$$

Using a Digital Elevation Model (DEM) of the earth's surface, $\hat{C}'$ can be projected from the location T_x, T_y, T_z until it intersects with a facet on the ground. If a DEM is not available, a flat earth may be assumed. In this case, the X, Y location of each projection in our world coordinate space could be determined by:

$$X_{loc} = T_x - \frac{\hat{C}'_x}{C'_z} H \tag{4.13}$$

$$Y_{loc} = T_y - \frac{\hat{C}'_y}{C'_z} H \tag{4.14}$$

Once the positions are known for each pixel on the ground, the VIS and NIR pixels which are closest to each other can be matched. It should be noted that this method will never be capable of perfectly registering the two datasets together because no two pixels image the same exact spatial region on the ground, but improvement can be made in the spatial registration, hopefully to within 1 pixel.

Chapter 5

MISI Registration

5.1 Registration Method

A geometric model of the MISI sensor was created using methods discussed in Chapter 4. Using this geometric model, and ray tracing through MISI's optical system, the angular difference between the ground-looking vectors of the VIS and NIR optical fibers at the scan mirror can be determined. Using equation 4.6, $\hat{C}$ vectors for both the VIS and NIR fibers were calculated as a function of the scan angle. We will call these vectors $\hat{C}_{VIS}$ and $\hat{C}_{NIR}$. Keep in mind, using this method, only the angular separation between fibers is used to characterize the misregistration. Recall that the MISI sensor has two optical fibers on the primary focal plane which are in the scan direction. However, due to the rotation of the scan mirror, the angular separation is not purely in the scan direction, and some component of the separation may be in the along-track direction. Since MISI collects data at a given rate, each pixel is imaged over a small range of scan angles. Since data is being captured over very small amounts of time, it can be estimated that each pixel is collected instantaneously at a given scan angle. At this point it is advantageous to express the $\hat{C}_{VIS}$ and $\hat{C}_{NIR}$ vectors in terms of their along and across track components. We can do this by writing them as:

$$\hat{C}_{VIS} = \hat{C}_{VIS-across}, \hat{C}_{VIS-along} \tag{5.1}$$

$$\hat{C}_{NIR} = \hat{C}_{NIR-across}, \hat{C}_{NIR-along} \tag{5.2}$$

Now the angle between the along track and across track components of the visible and near-infrared vectors can be found. An inner product is used to compute the angle. That is,

$$\gamma_{across} = \cos^{-1}\left(\frac{\hat{C}_{VIS-across} \cdot \hat{C}_{NIR-across}}{\left|\hat{C}_{VIS-across}\right|\left|\hat{C}_{NIR-across}\right|}\right) \tag{5.3}$$

$$\gamma_{along} = \cos^{-1}\left(\frac{\hat{C}_{VIS-along} \cdot \hat{C}_{NIR-along}}{\left|\hat{C}_{VIS-along}\right|\left|\hat{C}_{NIR-along}\right|}\right) \tag{5.4}$$

Plots of γ_{across} and γ_{along} for a ± 45 degree field of view are shown in Figure 5.1.

A raw MISI image samples a scene with 1500 pixels in the across-track direction. The across track FOV is $\frac{\pi}{2}$ radians. The across-track offset in pixels therefore may be given as

$$pixels_{across} = \frac{1500}{\frac{\pi}{2}}\gamma_{across} \tag{5.5}$$

Given the aircraft velocity and γ_{along}, the the number of pixels shifted in the along-track direction can be similarly found. The aircraft may be flown at a velocity such that oversampling occurs in the across-track direction. For instance, if the aircraft was flown such that $2x$ oversampling occurred in the across-track direction, twice as many pixels would need to be shifted in the along-track direction as would need to be shifted in the across-track direction for any given angle of separation. The along-track offset in pixels may be written as

$$pixels_{along} = \frac{1500}{\frac{\pi}{2}}\gamma_{across}S_{sampling} \tag{5.6}$$

where $S_{sampling}$ is the amount of oversampling in the across-track direction.

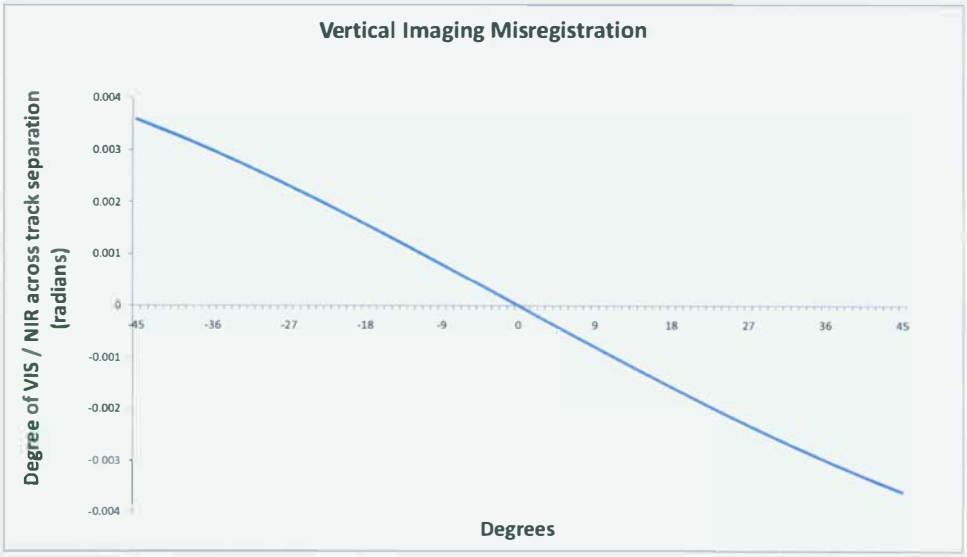
This is not necessarily a great method to use, since many assumptions are made. These assumptions include (1) a level flying aircraft, (2) no roll, pitch, or yaw, (3) constant velocity, and (4) flat terrain. A better co-registration approach would be one which incorporates some knowledge of the location data is being imaged on the ground, and using that information to perform a nearest-neighbor type of interpolation using the actual ground locations of each pixel to determine the nearest neighbor. For any image, the following would be done:

For each pixel in the image, compute X_{loc} and Y_{loc} for both the VIS and NIR set of bands. we'll call these $X_{VIS-loc_{i,j}}$, $Y_{VIS-loc_{i,j}}$ $X_{NIR-loc_{i,j}}$, and $Y_{NIR-loc_{i,j}}$ where i, j represents a pixel index. Then for each pixel in the the visible set of bands of the image, compute an Euclidean

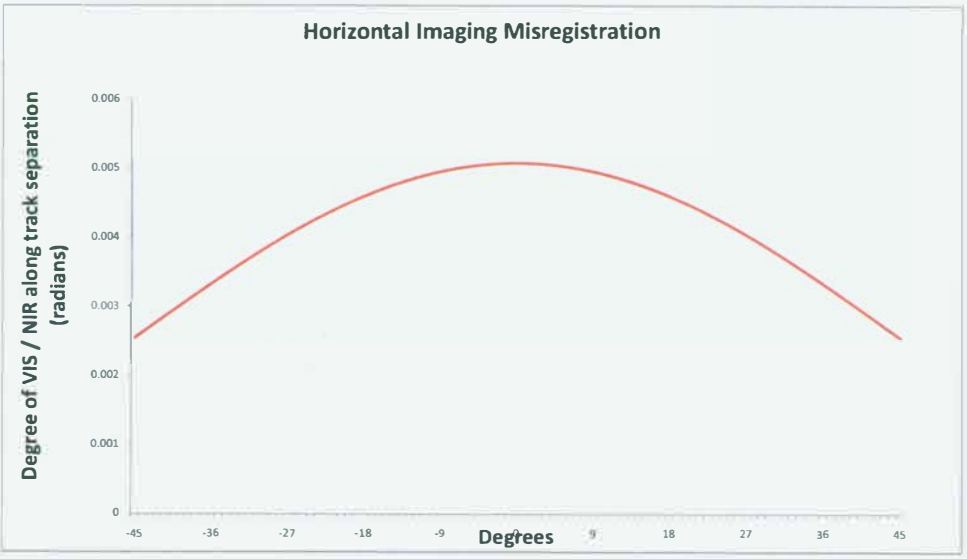
distance from $X_{VIS-loc_{i,j}}$, $Y_{VIS-loci,j}$ to every NIR pixel's ground coordinates, $X_{NIR-loc_{i,j}}$, $Y_{NIR-loci,j}$. This can be very computationally expensive if every pixel in the VIS set of bands is compared to every pixel in the NIR set of bands. Time complexities are usually given in Big-Oh notation to asymptotically describe the time taken for an algorithm to execute. For and $N \times N$ image the time complexity is $O(N^4)$. To cut down on computation time a windowed region may be used to search only local NIR pixels. A quick visual inspection can be performed to estimate the maximum pixel shift in the image. Creating a window several pixels larger than this shift will ensure all possible closest pixels will be included in the nearest neighbor search, and reduced the time complexity much closer to $O(N^2)$.

In MISI's particular case, finding γ_{along} and γ_{across} cannot be achieved with too much accuracy due to noisy Digital Measurement Unit (DMU) measurements and the fact that GPS data is collected at discrete time intervals which are quite large relative to pixel capture rates. GPS data is collected about every 20-30 lines or so. A linear interpolation is used to estimate the aircraft's position for every pixel between GPS measurements. A mean is computed for each line of roll, pitch and yaw data (heading for some datasets) for each line in an image. This means in MISI's case, we assume the plane has the same position and orientation for an entire line. Although accurate data is not present for each pixel in an image, this method does show improvement over the method using only angular offsets and no aircraft data.

The methods discussed in this chapter have been implemented. Results using only the sensor model, as well as results using the sensor model along with aircraft orientation data will be shown in chapter 7.



(a) Amount of vertical misregistration (γ_{along})



(b) Amount of horizontal misregistration (γ_{across})

Figure 5.1: Amount of horizontal and vertical misregistration in an scan line for a plane flying straight and level. The X-axis represents the angle of the scan mirror from -45 to +45 degrees. The Y-axis represents the amount of spatial misregistration in radians.

5.2 MISI Registration Error Analysis

There are many sources which contribute to error in MISI's registration process. Mirror wobble, noisy DMU measurements, and interpolation between sparse GPS measurements using the assumption of constant aircraft velocity are all examples of error sources that affect the accuracy of MISI's registration. Because a nearest neighbor interpolation is used in the registration process, a rigorous error analysis including all error sources is beyond the scope of our goals. A simple error analysis is performed below using the geometric model developed Chapter 4. For the images used in this thesis, MISI was flown at about 1000m. MISI's parameters at a flying height of 1000m are shown in Table 5.1. It is seen in Figure 5.1 at NADIR, there is no

Table 5.1: MISI Parameters at 1000m

Fixed Parameters	
Minimum Flying Height	1 km
Minimum Ground Speed of Aircraft	110 mph (49 m/sec)
IFOV	2.54 milliradian
VIS NIR fiber separation	5.08 milliradian
Pixels per Line	1500
FOV	90 deg
Derived Parameters	
Minimum GIFOV	2.54 m
Scan Rate at 1 km	27.3 Hz

along-track misregistration, but there is a roughly 5 milliradian across-track misregistration. Using the number of pixels per line and the FOV in table 5.1, we compute a pixel-to-pixel scan angle of 1.05 milliradian. The pixel-to-pixel scan angle is the amount the scan mirror rotates between 2 successive imaged pixels. At a flying height of 1000m this would correspond to a 1.05 meter misregistration on the ground. Also recall from Figure 5.1 the across track misregistration near the center of the image is roughly 5 milliradians, while there is almost no along track misregistration. This means registration near the center of the image can be achieved by translating a VIS image roughly 5 pixels relative to a NIR image in the across track direction. The maximum along track misregistration is roughly 3.5 milliradians, which corresponds to 3.5 meters on the ground at a flying height of 1000m. Given a constant aircraft velocity of $49\frac{m}{s}$ and a scan rate of 27.3 Hz, the along track sampling is computed to be 1.79 meters. Using the previously computed conversion factor of 1.05 milliradians per pixel at a 1000m flying altitude, the maximum along track misregistration is 1.7 pixels, or 1.7 lines. These computed along and across track misregistration amounts are consistent with what has been observed in raw MISI

imagery.

If the error sources considered include aircraft position and orientation, we can determine the changes needed to produce an error of 1 pixel after registration has been applied. Using the numbers in Table 5.1 the amount of time elapsed between samples can be computed. For simplicity, we'll assume a maximum across track misregistration of 5 pixels, and a maximum along track misregistration of 2 lines. The amount of time between 2 scan lines (assumed maximum along track misregistration) is computed below.

$$t_{along} = 2l\frac{1s}{27.3l} = 0.073s \tag{5.7}$$

where l denotes 1 line. The amount of time between 5 samples (across track misregistration) is computed below

$$t_{across} = 5p\frac{1l}{1500p}\frac{1s}{27.3l} = 1.22 \cdot 10^{-4}s \tag{5.8}$$

where p denotes 1 pixel. Using these times, we can compute the aircraft acceleration, or the angular velocity required to cause a shift of 1 pixel. For simplicity, we'll assume a pixel to pixel misregistration of 1 meter in the across track direction and line to line misregistration of 1.8 meters in the along track direction (due to oversampling, recall the along track sampling is 1.79 meters). To induce a shift of 1 pixel by translation of the aircraft alone, the aircraft would need to move $1m$ in across track or $1.8m$ in the along track direction within the times given in equations 5.7 or 5.8. Assuming an initial constant aircraft velocity, using equation 5.9, we can compute the accelerations needed to cause this shift.

$$\Delta x = \frac{1}{2}at^2 \tag{5.9}$$

rearranging we get

$$a = \frac{2\Delta x}{t^2} \tag{5.10}$$

Using values of $1.8m$ for x and $0.073s$ for t in the case of along-track misregistration, an acceleration of roughly $671m/s^2$, or $68.4g$ is required to produce a shift of 1 pixel, where g is earth's gravitational constant. Such a large acceleration shows aircraft acceleration can account only for shifts much smaller than 1 pixel. It is also safe to assume that the effects of aircraft translation are negligible when considering only across track misregistration where the

time between scanned pixels is much shorter. This means changes in aircraft position have a negligible effect on misregistration errors..

Changes in aircraft orientation are also an error source. For an aircraft flying perfectly level during image aquisition, there will be no error due to orientation. The aircraft orientation will change during flight, and again, we compute the angular velocity required to cause a shift of 1 pixel between spectral bands. The amount of roll required to cause a shift of 1 pixel in the across track direction is given below.

$$\frac{d(\omega,\phi)_{best}}{dt} = \frac{1 milliradian}{1.4 \cdot 10^{-4} s} = 7.14 \frac{rad}{s} = 409 \frac{deg}{s} \tag{5.11}$$

$$\frac{d(\omega,\phi)_{worst}}{dt} = \frac{1.8 milliradians}{0.073 s} = 24.6 \frac{mrad}{s} = 1.4 \frac{deg}{s} \tag{5.12}$$

where (ω, ϕ) represents the roll of the aircraft in equation 5.11 or pitch of the aircraft in equation 5.12. Equation 5.11 represents the best case scenario, applying registration near the center of the image, where across track misregistration is dominant, and along track misregistration is zero. Equation 5.12 represents the worst case scenario, applying registration near the edges of the image where along track misregistration is abundant, and more time has elapsed between the VIS and NIR pixels to be paired together. It can clearly be seen here that uncertainties in aircraft orientation can cause significant errors in the registration process. Error sources are much more significant in the along track direction than the across track, so it is difficult to register pixels far from the center of an image.

The error analysis discussed in the above paragraphs account for aircraft dynamics. Other sources of error inherent to the instrument itself exist as well. Non-uniformities in the fold and scan mirrors, scan mirror wobble, uncertainty in the position of the primary focal plane and uncertainty in the position of the fibers feeding the VIS and NIR spectrometers are all error sources which can effect the registration process. Laboratory measurements can be conducted to better measure these things. If a laboratory experiment is performed to characterize the misregistration itself, the uncertainties in many of these parameters would be accounted for in that measurement, and a better model of the sensor could be constructed.

Chapter 6

Effects of Misregistration on Image Data

In previous chapters it was noted that misregistration between sets of spectral bands would affect the performance of target detection algorithms. This is because target pixels in the misregistered data sets will not yield an accurate representation of the target spectrum. It has been shown using thematic mapper data that the use of misregistered data has negative effects on land and crop classification (27). In that particular study, classification changed significantly, up to 10% for band-to-band misregistrations as little as 0.3 pixels. Larger amounts of misregistration up to 3 pixels resulted in a 30% change in classification in some cases. It is interesting to notice that large changes occur despite the data being imaged over uniform regions. It would seem that since neighboring pixels are spectrally similar, the misregistration should not have much impact on classification. Nonetheless, the results indicate that good registration is critical to good classification performance.

A technical report written for the Jet Propulsion Laboratory states that for land classification sensor reaching radiances should come from the same spot on the ground at the 95% level when using multi or hyperspectral sensors. (28)Larger amounts of misregistration will result in larger errors in retrieved radiance and derived reflectance values. Spectra with these induced errors will compromise analysis algorithms because many observed spectra will not be physically realizable. This will result in misclassification of materials, for instance, because the observed spectra will not be represented by any spectral library or any reflectance curves that

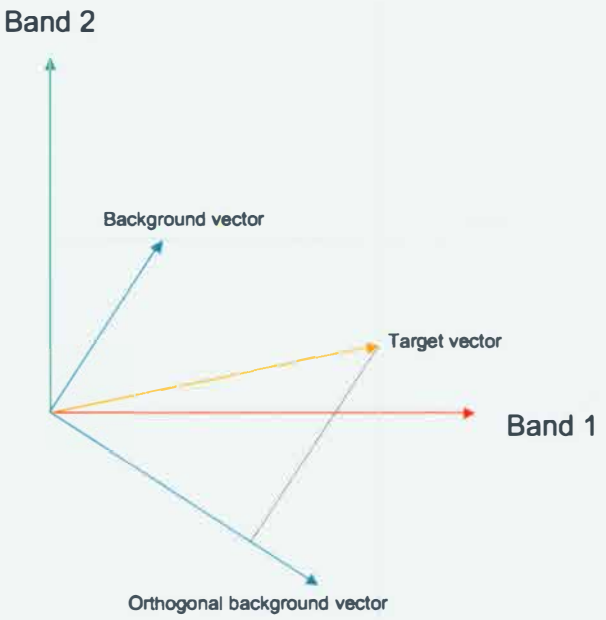


Figure 6.1: A two band case showing a target vector projected onto a space orthogonal to the background.

might be in a geological database.

Although band-to-band misregistration can cause large errors in classification routines, little information is available on its effects on target detection. Geometric algorithms use background endmembers to characterize the background. Operations are performed to characterize a space orthogonal to the background, and a vector representing the spectrum of a pixel can be projected onto this space to determines how target-like a pixel is. A simple illustration of a geometric target detection operation is illustrated in Figure 6.1 for a simple case using 2 spectral bands.

Using the simple case in Figure 6.1, the effects on target detection performance can be demonstrated if band 2 is shifted spatially relative to band 1. First consider an image in which a target occupies exactly 1 pixel. In the 2 band case, the target pixel should resemble the target vector, since the target occupies the entire pixel and there is no background mixing in the pixel. Next consider the same case, only now band 2 is spatially shifted relative to band 1. Since no change was applied to band 1, the magnitude of the target pixel vector in band 1 will remain unchanged. The magnitude of the target pixel vector in band 2, however, will be changed because of the spatial shift. This effect is illustrated in Figure 6.2. Since the target in this example occupies exactly 1 pixel, any spatial shift will cause a fractional mixing between background and target signatures. These types of linear mixture can be expressed in equation

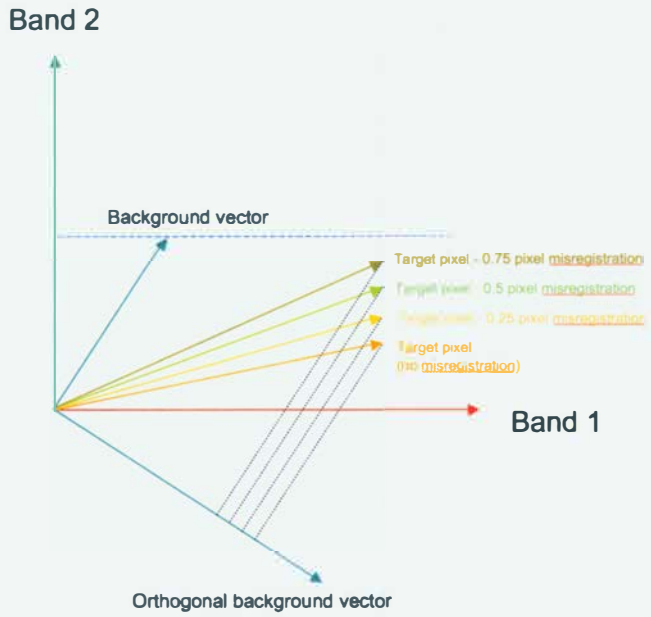


Figure 6.2: 2 band case showing a target vector projected onto a space orthogonal to the background. The target vectors have been shifted by 0.25, 0.5 and 0.75 pixels. Notice the magnitude in band 1 does not change for any of the target pixels, but the magnitude in band 2 changes based on the amount of background mixed in the target pixel.

6.1 where subscripts 1 and 2 denote the band number, and α denotes the fractional amount of target which occupies the pixel. Since many sensors employ more than 2 spectral bands, this equation may be generalized to accommodate such sensors.

$$\begin{aligned} x_1 &= d_1 \\ x_2 &= \alpha d_2 + (1-\alpha) b_2 \end{aligned} \tag{6.1}$$

This can be seen in equation 6.2 where $bands1$ represents the set of bands with target present and $bands2$ represents the set of bands which is spatially shifted relative to $bands1$. In the case of MISI, $bands1$ could represent the VIS channels and $bands2$ could represent the NIR channels.

Stochastic algorithms are also affected by misregistration, although because of the more complex math involved in these types of algorithms, the effects are much harder to illustrate. This forces us to rely on trends in the data, and comparisons between stochastic and geometric methods to make conclusions on the performance of each target detection algorithm.

$$\begin{aligned} x_{bands1} &= d_{bands1} \\ x_{bands2} &= \alpha d_{bands2} + (1-\alpha) b_{bands2} \end{aligned} \tag{6.2}$$

It is important to note that the phenomena expressed in equation 6.2 occurs not only for target pixels, but for every pixel in the scene. This means that for every pixel in the scene being observed in *bands*1 will have some spectral mixing with adjacent pixels in *bands*2. Since this strange type of spectral mixing between sets of spectral bands occurs for every pixel in the scene, background characterization is affected, especially when observing pixels at natural borders in the scene. For instance, if we consider a scene which contains grass pixels and asphalt road pixels as background endmembers, spectral mixing in one set of spectral bands will occur where the 2 background endmembers meet. Pixels at this boundary will contain a certain linear mixture of grass and road in *bands*1, but a different linear mixture of grass and road in *bands*2. For this reason, it has been shown that change detection using misregistered multispectral data produces false changes which are equally distributed along the edges in the imagery (29).

6.1 Simulation Experiment

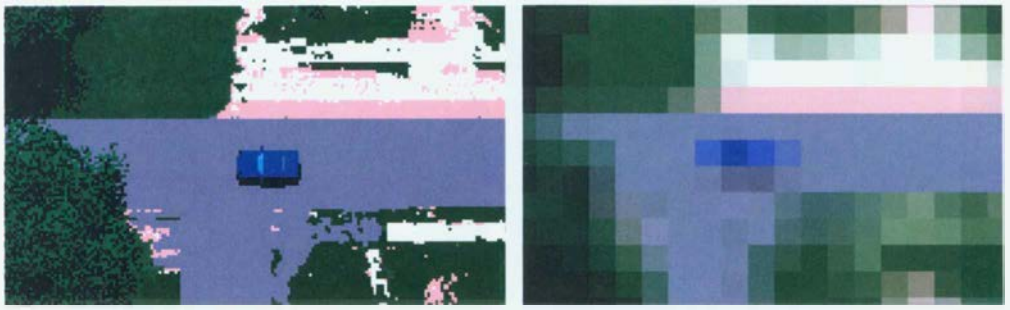
6.1.1 Scene of interest

To analyze the effects of misregistration on hyperspectral target detection, simulated imagery was used. Simulations can be a great tool to compare the performance of hyperspectral target detection algorithms under varying levels of misregistration because simulated imagery produced using a scene model provides a great level of control. In this particular study, 140 vehicles were placed in a scene and DIRSIG was used to capture images over the area of interest using a simulated hyperspectral sensor. Tile 1 of Megascene 1 (30)was chosen as the scene which can be seen in Figure 6.3. The material map was slightly modified for to allow vehicles to be parked on 3 different road surfaces as well as 2 grass fields. Red Toyota sedans can be seen in the images in Figure 6.3.

Figure 6.3: Tile 1 of megascene. Red Toyota sedans can be seen on the roads and in the grass fields in the image.

6.1.2 Creating misregistered images

The simulated image collection was performed using a hyperspectral sensor with 73 evenly spaced bands covering the 0.4-1.12 μm region of the electromagnetic spectrum. A spatial shift between the VIS and NIR channels was chosen to resemble the shift between MISI's VIS and NIR bands. To simulate the spatial shift, the DIRSIG imagery had to be oversampled. The resulting image with misregistration between the VIS and NIR bands was created in the following manner: First oversampled imagery of the scene was generated using DIRSIG. The scene was generated using ideal sampling, as a sensor point spread function (PSF) would be



(a) Subsection of 10x oversampled image. A vehicle model with blue Ford focus paint applied can be seen.

(b) The same subsection after the image has been degraded.

Figure 6.4: Oversampled and degraded imagery. Figure (a) shows a blue Ford focus in the oversampled imagery. Figure (b) shows the same vehicle after the image has been degraded by the simulated sensor PSF.

applied to the oversampled imagery later. The scene was 10x oversampled at 4000 x 4000 pixels to allow spatial shifts in 0.1 pixel increments. A shift of 1 pixel in the oversampled imagery corresponds to shift of 0.1 pixel in the resulting degraded imagery. After the spatial shift was applied to the oversampled imagery, it was degraded using a circularly symmetric gaussian to simulate the point spread function of a sensor. This was chosen because a circularly symmetric gaussian gives a good representation of a sensor response (31). The gaussian used was slightly wider than the area being degraded to allow some blending of adjacent pixels, as this realistically represents a sensor PSF. This process was performed for each oversampled image for the case of 0, 0.1, 0.2, 0.3, 0.4, and 0.5 pixel misregistration. The resulting degraded, misregistered imagery had dimensions of 398 columns by 397 rows. A section of the oversampled image containing a vehicle, as well as the resulting degraded image can be seen in Figure 6.4. The vehicle seen in Figure 6.4 is a sedan with blue Ford focus paint applied to the body of the car.

6.2 Preprocessing images

6.2.1 DIRSIG images

After the degraded misregistered images were created, noise was added. Signal to noise ratios (SNRs) of 20, 100 and 500 used. The SNR of 20 was chosen to imitate MISI's estimated noise characteristics, while 100 and 500 were chosen to simulate higher end sensors and to determine how various amounts of noise affect target detection performance when using misregistered

data. Noise was added to the images by creating a gaussian noise image with dimensions equal to dimensions of the degraded images. The noise images created had a mean of zero and a standard deviation of 1. SNR is given as

$$SNR = \frac{S}{N} \tag{6.3}$$

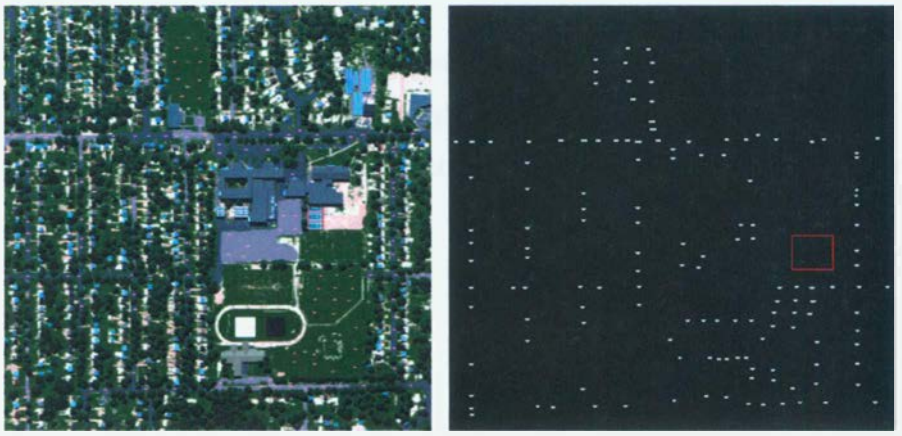
where S is the signal and N is the standard deviation of the noise. Since the noise image had a standard deviation of 1, each pixel in the scene was multiplied by the noise image and divided by the corresponding SNR to construct a new noise image which, when added to the noise free image would produce a noisy image with the correct SNR. To ensure noise was added correctly, a random band in the noisy imagery was chosen and the standard deviation was calculated using all the pixels in that band. The calculations showed the computed noise statistics matched closely with the SNR specified for the image, thus the routine written to add noise to the image was implemented correctly.

After noise was added to the images ELM atmospheric compensation was applied. Notice there are two large calibration panels of known reflectance inside the track area in Figure 6.3. Atmospheric compensation was the last preprocessing step before target detection algorithms were applied.

To prepare the images for target detection experiments, the background was characterized. A target mask was created to remove vehicles and produce a target free scene. This target free scene was used to compute a covariance matrix for background classification using the statistical CEM and ACE algorithms. The target free scene was also used to perform k-means clustering to extract endmembers for the OSP and ASD geometric algorithms. Six endmembers were used based on the approximate number of dominant materials throughout the scene.

6.2.2 Target detection performance

To assess target detection performance, 140 identical vehicles were placed in the scene (repeated 4 times for the different vehicles in Table 6.1), and a target map was created to classify pixels as either target or non-target in the resulting degraded imagery. The degraded imagery along with a vehicle location truth map can be seen in Figure 6.5. The vehicle truth map was used to calculate the probability of detection as a function of false alarm rate. This information was used to construct receiving operator characteristic (ROC) curves.



(a) Degraded Tile1 scene and vehicle location truth

(b) Truth locations of vehicles.

Figure 6.5: Degraded imagery and vehicle location truth

Before target detection algorithms were applied to the misregistered degraded imagery, the algorithms were applied to the imagery using the VIS and NIR channels separately. This was done to determine whether using the entire set of spectral bands results in better performance for the algorithms, and in the cases where using the entire VIS/NIR spectral range does give improvement, how much misregistration between bands is tolerable. A listing all vehicles used and experiments performed can be seen in Table 6.1

Table 6.1: Misregistration experiments

Vehicle	SNR	Amount of misregistration (pixels)
Red Toyota	SNR=20, SNR=100, SNR=500	0, 0.1, 0.2, 0.3, 0.4, 0.5
Blue Ford Focus	SNR=20, SNR=100, SNR=500	0, 0.1, 0.2, 0.3, 0.4, 0.5
Green BMW	SNR=20, SNR=100, SNR=500	0, 0.1, 0.2, 0.3, 0.4, 0.5
White Saturn Vue	SNR=20, SNR=100, SNR=500	0, 0.1, 0.2, 0.3, 0.4, 0.5

A complete listing of results is given in section 7.2.

The data used for all DIRSIG simulated experiments was ELM atmospheric compensated data. An ELM was performed using calibration panels of known reflectance inside the track area in the scene. Weak water absorption features exist in the spectral range that was used in the experiment, but were not strong enough to affect retrieved spectral signatures. Because of this, no bad bands list was needed for the simulation experiments.

Target detection algorithms discussed in Chapter 3 were used to detect vehicles. Vehicle reflectance signatures were obtained from a spectral library. These signatures were used to

apply a spectrum to the vehicles in the scene and were also used as the known target signature in the target detection algorithms used. The algorithms used were SAM, OSP, ASD, CEM and ACE. Data was mean subtracted when computing covariance matrices for the CEM and ACE algorithms.

Chapter 7

Results

This chapter contains two sections, one for MISI registration and another for target detection results using both real imagery and simulated DIRSIG imagery. The section on MISI will describe the improvement in registration between the VIS and NIR channels after applying the registration techniques discussed in Chapter 5. The section on target detection results will discuss target detection algorithms applied to MISI imagery, although the main focus is target detection performance on simulated misregistered DIRSIG imagery.

7.1 Registration Results

Two variations of co-registration have been implemented. One method uses the geometric model to determine the angular offset between bands as a function of the scan angle. This method has advantages in that aircraft orientation parameters are not required when dealing with across-track registration. However it has shortcomings when dealing with along-track registration. This becomes obvious when thinking of the along-track angular offset. Since data in the along-track is imaged as the aircraft moves forward, the difference in aircraft position between collected lines of data must be known in order to accurately adjust for the along-track angular offset. For the preliminary results presented, measurements from the onboard DMU were very noisy. Therefore they were not incorporated into the registration process in the first preliminary step.

Each figure in this section is an image of a portion of RIT collected by MISI. Figure 7.1 is an imaged section of the RIT campus under observation. All subsequent images are subsections

Figure 7.1: Full FOV region of a MISI image.

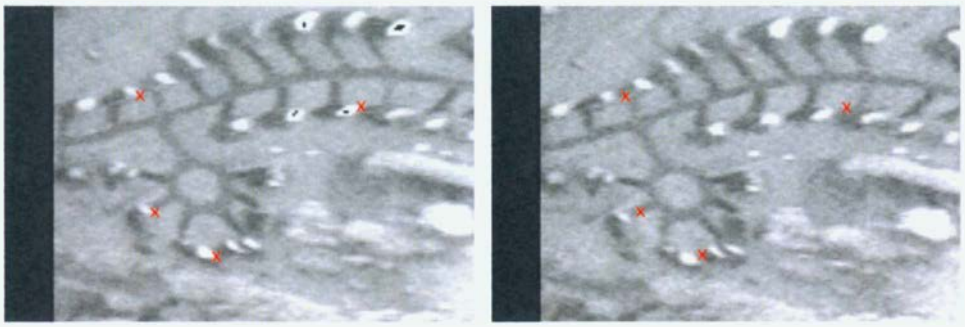
of this image, and are included to demonstrate the misregistration artifacts before and after co-registration has been applied to the images. The subset images are located at the left, center and right side of Figure 7.1.

7.1.1 Registration using sensor model only

Figures 7.2 through 7.4 show the last band from the VIS spectrometer, and the first band from the NIR spectrometer before any co-registration has been applied. These images contain the same set of pixels, while markers have been placed at the same pixel index to represent the misregistration between bands. The fact that no registration has been attempted on these images should be apparent as the markers vary in ground location by up to several pixels.

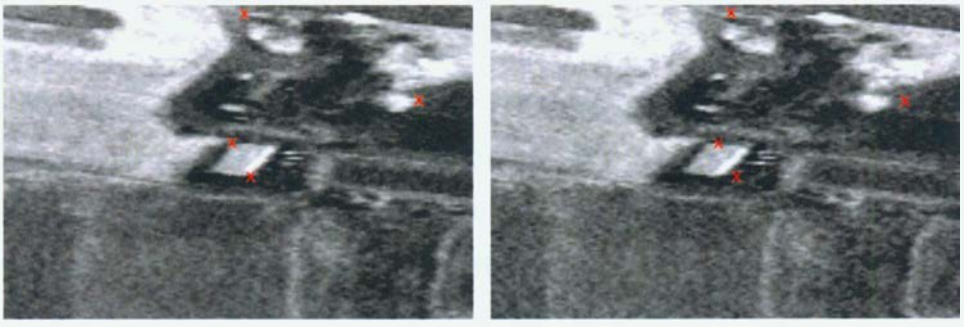
Images 7.5 through 7.7 show the last band from the VIS spectrometer, and the corresponding set of pixels from the NIR spectrometer after co-registration has been applied using the angular offsets from the geometric model. Aircraft parameters were estimated here using the following assumptions: 1.5 oversampling, flat and level aircraft flight, constant aircraft velocity and constant height. These assumptions are not very good, as lightweight aircraft are affected by turbulence and wind. The pixels to be registered to each other are imaged at time intervals that are relatively close to one another, and some of the assumptions we've made about the aircraft dynamics may be acceptable over short periods of time.

No real aircraft parameters were used, and this first step served as a check to ensure the model provided an approximate solution to the registration problem. Notice that while some features between images match up very well, other features do not. While a very clear misregistraiton pattern was apparent in the raw data, there seems to be little pattern to the misregistration here. Some features in the co-registered NIR image are shifted upward and to the right, while others are shifted downward and to the left relative to the VIS image. There are several reasons for this:



(a) Visible band, left side of image (b) Near infrared band, left side of image

Figure 7.2: Spatial misregistration between VIS and NIR bands at the left side of the image. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.



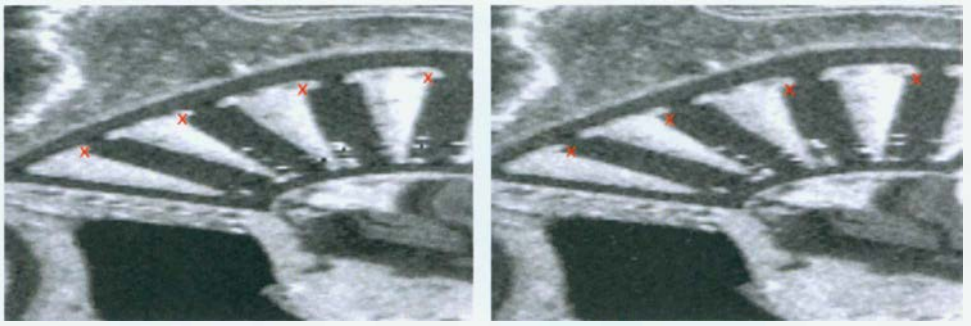
(a) Visible band, center of image (b) Near infrared band, center of image

Figure 7.3: Spatial misregistration between VIS and NIR bands at the center of the image. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

1) No aircraft parameters were taken into account using this model - Changes in X,Y position, height, roll, pitch, and yaw will determine where data is being imaged on the ground for each pixel and they are not taken into account.

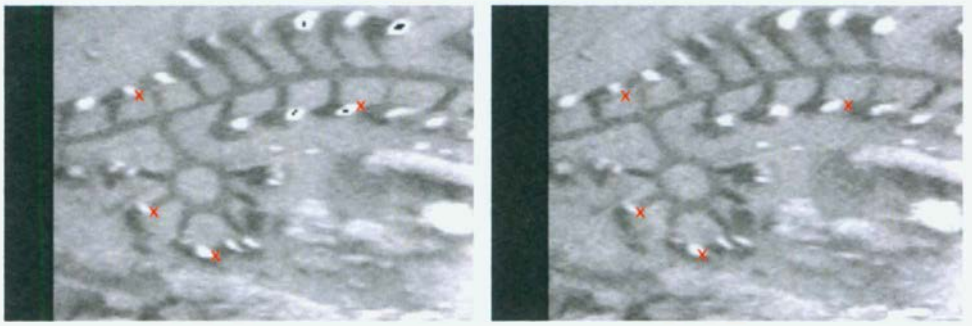
2) Aircraft Velocity is also assumed constant, so along track misregistration features will become stretched and compressed as the aircraft changes speed.

3) The angle of separation between the ground-pointing vectors increases as a function of scan angle in both the across-track and the along-track. Since the angle of separation increases, as well as the distance to the ground increases at larger scan angles, the error in co-registration may also increase.



(a) Visible band, right side of image (b) Near infrared band, right side of image

Figure 7.4: Spatial misregistration between VIS and NIR bands at the right side of the image. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

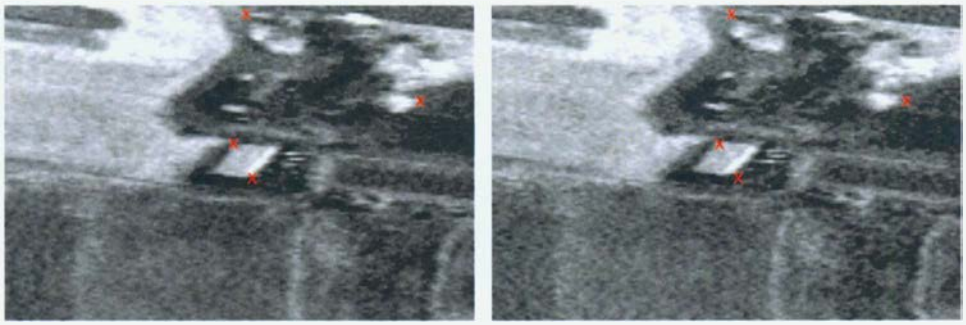


(a) Visible band, left side of image (b) Near infrared band, left side of image

Figure 7.5: Spatial misregistration between VIS and NIR bands at the left side of the image after co-registration has been applied to the images using the angular offset approach. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

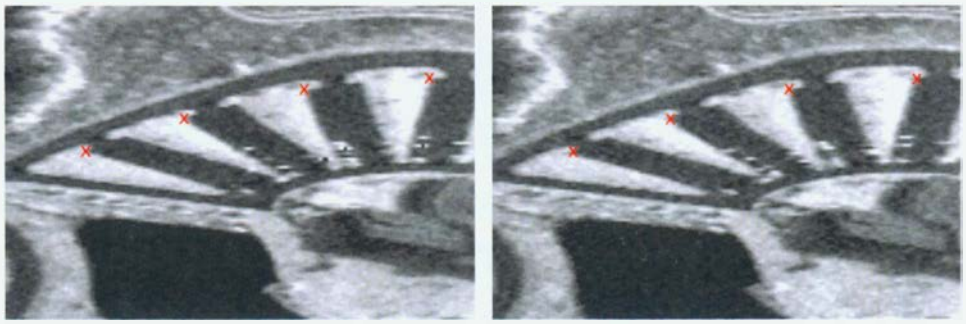
7.1.2 Registration using Aircraft Paramters

The second method utilizing ground location coordinates for each pixel then using a nearest neighbor interpolation was also implemented. Two cases were used, one using all aircraft orientation parameters, and another case using only the aircraft GPS coordinates. A visual inspection revealed no noticeable difference between the case where roll, pitch and yaw were used and when they were not. Although no noticeable change was seen for these two cases, using GPS data does seem to show improvement over the alternative method of using the sensor model only. Results can be seen below, although it is difficult to demonstrate the



(a) Visible band, center of image (b) Near infrared band, center of image

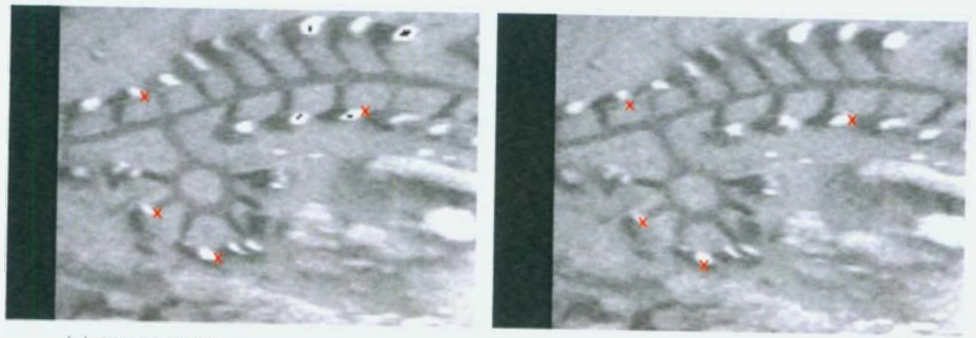
Figure 7.6: Spatial misregistration between VIS and NIR bands at the center of the image after co-registration has been applied using the angular offset approach. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.



(a) Visible band, right side of image (b) Near infrared band, right side of image

Figure 7.7: Spatial misregistration between VIS and NIR bands at the right side of the image after co-registration has been applied using the angular offset approach. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

quality of registration without linking displays and toggling between images. These sets of images can be seen in Figures 7.8 through 7.10. It is also difficult to quantify the results, as accurate measurements of the sensor have not been made. Even with an accurate sensor model, parameters such as mirror wobble, or mirror distortions due to centrifugal forces on the mirror during rotation will be difficult to estimate.

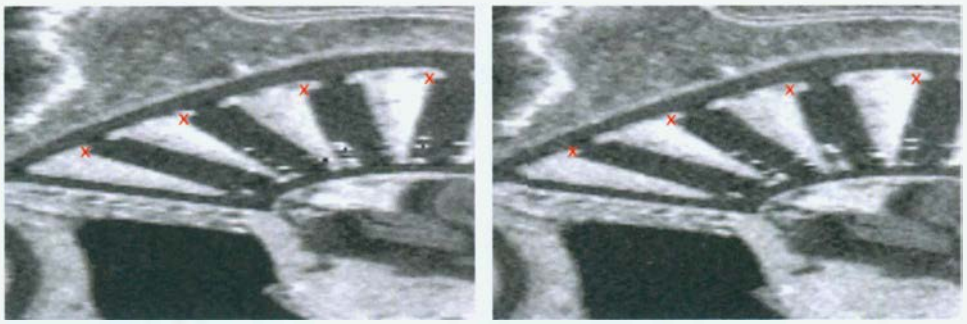


(a) Visible band, left side of image (b) Near infrared band, left side of image

Figure 7.8: Spatial misregistration between VIS and NIR bands at the left side of the image after co-registration using the nearest neighbor approach has been applied to the images. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

(a) Visible band, center of image (b) Near infrared band, center of image

Figure 7.9: Spatial misregistration between VIS and NIR bands at the center of the image after co-registration has been applied using the nearest neighbor approach. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.



(a) Visible band, right side of image (b) Near infrared band, right side of image

Figure 7.10: Spatial misregistration between VIS and NIR bands at the right side of the image after co-registration has been applied using the nearest neighbor approach. The red X's in the images denote the same pixel index location. Image (a) shows the last set of bands imaged by the visible spectrometer and (b) shows the first set of bands imaged by the infrared spectrometer.

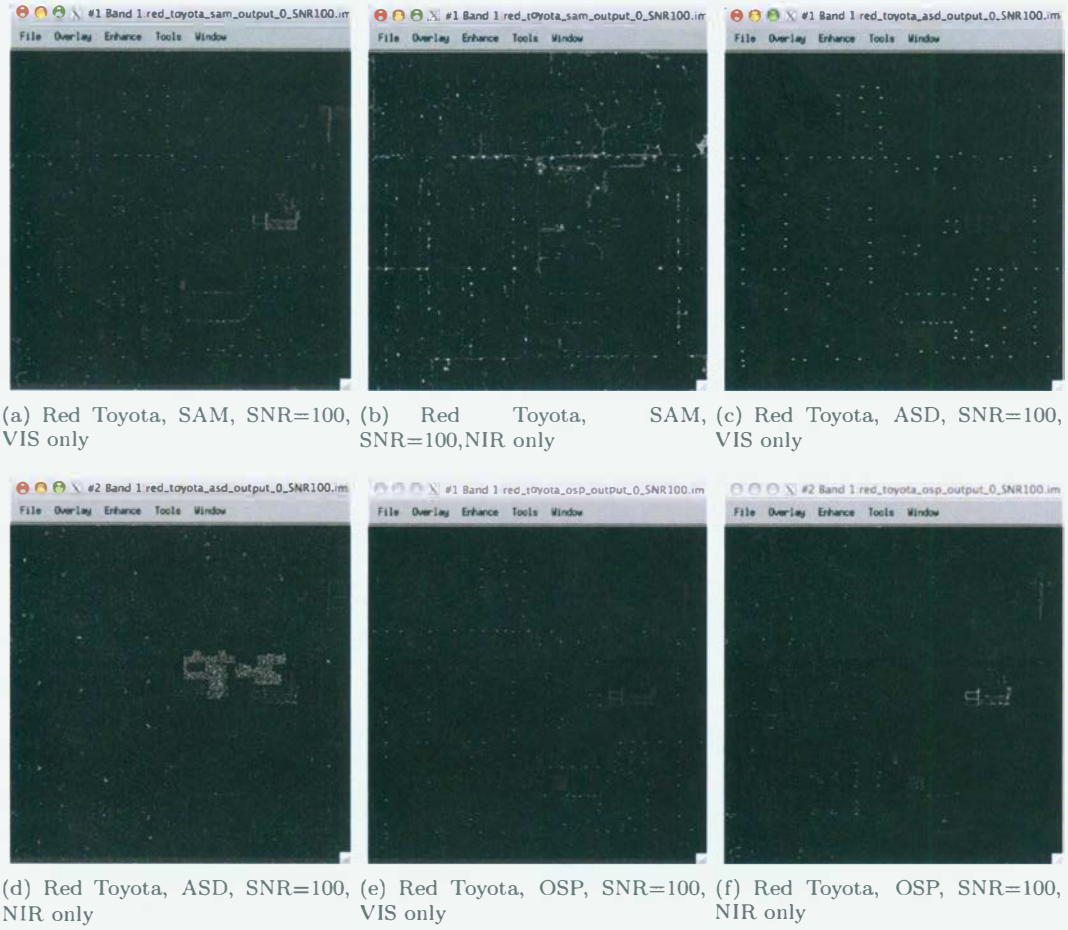
7.2 DIRSIG Target Detection Results

Results are shown here for the simulated misregistered target detection experiment described in Chapter 6. While many detection images were created during the process, only the detection maps from imagery with a SNR of 100 will be shown here. All images in this section are presented to give the reader a qualitative view the effects spatial misregistration have on hyperspectral target detection. The reader should look at the ROC curves in this section to get a more quantitative understanding of misregistration effects on target detection performance.

Some detection maps for the scene with red Toyotas parked throughout can be seen in Figures 7.11 and 7.12. Figure 7.11 shows detection maps for SAM, and the geometric target detection algorithms ASD, and OSP using the VIS and NIR channels separately. Figure 7.12 shows detection maps for the stochastic target detection algorithms CEM and ACE. This step was performed for two reasons: to determine if some detection algorithms perform better using VIS and NIR bands separately, and to compare these detection results against results using the full, misregistered spectrum. The geometric algorithms ASD and OSP have been grouped together since they both rely on background endmembers for background classification. The stochastic algorithms CEM and ACE have been grouped together for the same reason; they both rely on the scene-wide covariance for background classification. Detection maps were made for each scene, but only the detection maps for scenes with a SNR of 100 are shown in this document for the sake of eliminating redundant figures and saving space. The same type of detection maps, using only the VIS and NIR bands, can be seen for the blue Ford Focus in Figures 7.13 and 7.14, the green BMW in Figures 7.15 and 7.16, and the white Saturn Vue in Figures 7.17 and 7.18.

The same algorithms were applied to the degraded, misregistered images using all available bands. Detection results for images with red Toyotas as targets can be seen in Figures 7.19, 7.20, 7.21, 7.22, and 7.23, the blue Ford Focus results are shown in Figures 7.24, 7.25, 7.26, 7.27, 7.28, the green BMW results in Figures 7.32, 7.33, 7.34, 7.35, 7.36, and the white Saturn Vue results in Figures 7.37, 7.38, 7.39, 7.40, 7.41. Each image in the sequence has a progressively higher level of misregistration, from 0 to 0.5 pixel. Again all these images are detection results from the algorithms run on the imagery with a SNR of 100. All the images have been stretched to display roughly the top 5% of pixels.

Image (b) in Figure 6.5 is an example of an ideal target detection result. This is a truth



(a) Red Toyota, SAM, SNR=100, VIS only

(b) Red Toyota, SAM, SNR=100,NIR only

(c) Red Toyota, ASD, SNR=100, VIS only

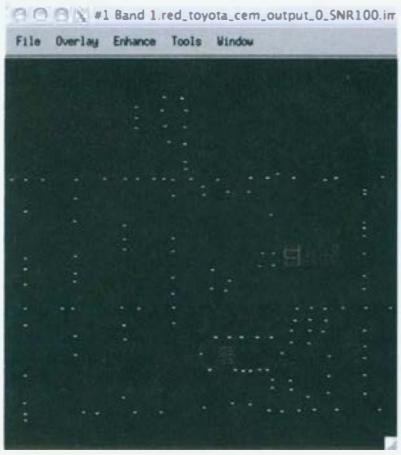
(d) Red Toyota, ASD, SNR=100, NIR only

(e) Red Toyota, OSP, SNR=100, VIS only

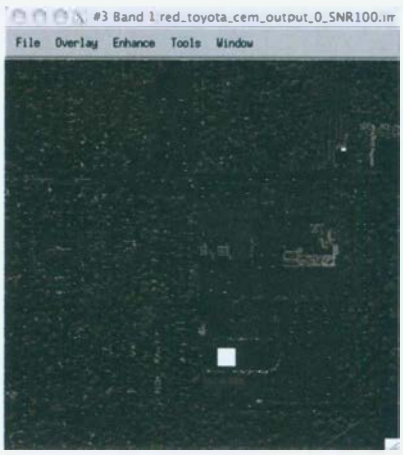
(f) Red Toyota, OSP, SNR=100, NIR only

Figure 7.11: detection images for the scene with red Toyotas parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the inverted output of SAM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ASD using only the VIS and NIR channels respectively, and Figures (e) and (f) show the output of OSP in the same manner. The histogram of these output images were all stretched to display roughly the highest 5% of detections.

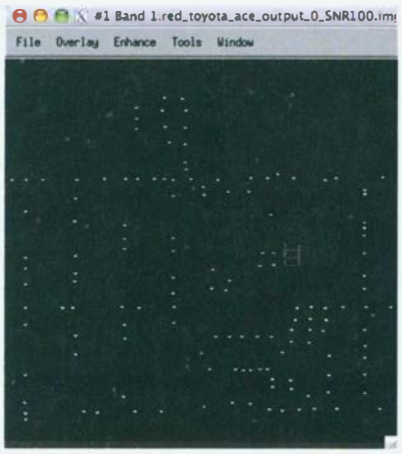
image, where bright pixels represent targets and dark pixels represent non targets. The image shows 100% detection with zero false alarms. Actual detection images will look similar to this, but background pixels will have some brightness. If background pixels are brighter than target pixels, this means there is a false alarm at that pixel. Good target detection results will look similar to the truth image in Figure 6.5, while bad detection results will contain brighter background pixels, and darker target pixels.



(a) Red Toyota, CEM, SNR=100, VIS only



(b) Red Toyota, CEM, SNR=100, NIR only



(c) Red Toyota, ACE, SNR=100, VIS only

(d) Red Toyota, ACE, SNR=100, NIR only

Figure 7.12: detection images for the scene with red Toyotas parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the output of CEM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ACE using only the VIS and NIR channels respectively. The histogram of these output images were all stretched to display roughly the highest 5% of detections.

(a) Blue Ford Focus, SAM, SNR=100, VIS only

(b) Blue Ford Focus, SAM, SNR=100,NIR only

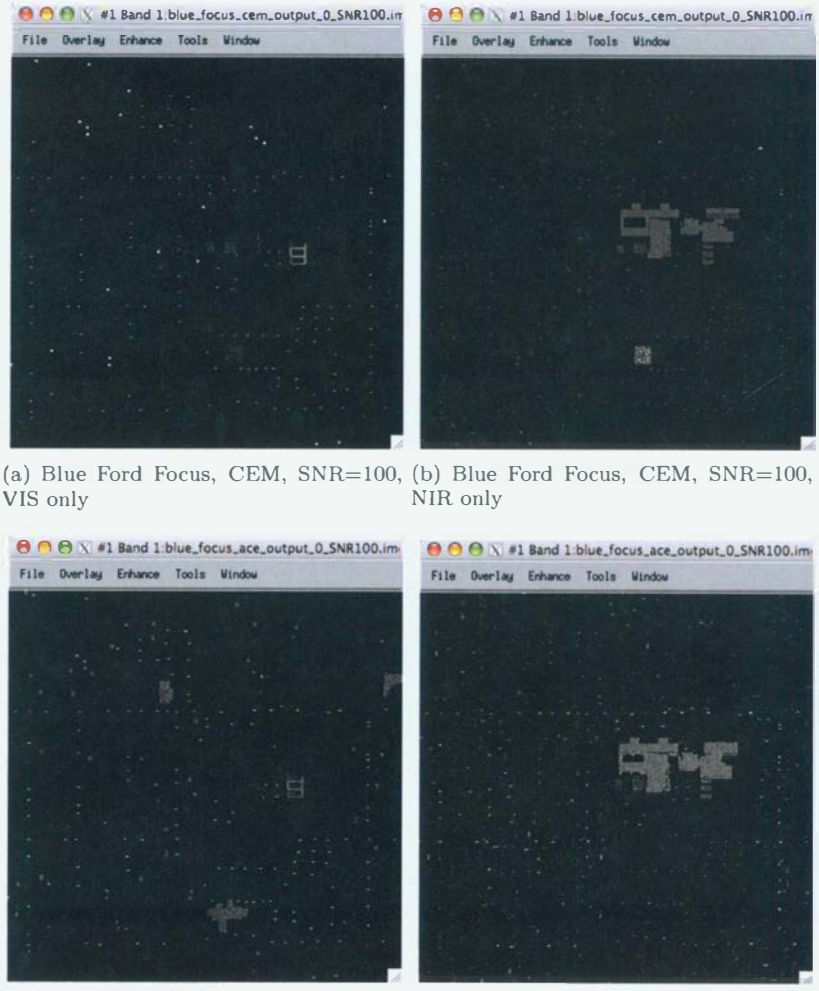
(c) Blue Ford Focus, ASD, SNR=100, VIS only

(d) Blue Ford Focus, ASD, SNR=100, NIR only

(e) Blue Ford Focus, OSP, SNR=100, VIS only

(f) Blue Ford Focus, OSP, SNR=100, NIR only

Figure 7.13: detection images for the scene with blue Ford Focuses parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the inverted output of SAM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ASD using only the VIS and NIR channels respectively and Figures (e) and (f) show the output of OSP in the same manner. The histogram of these output images were all stretched to display roughly the highest 5% of detections.



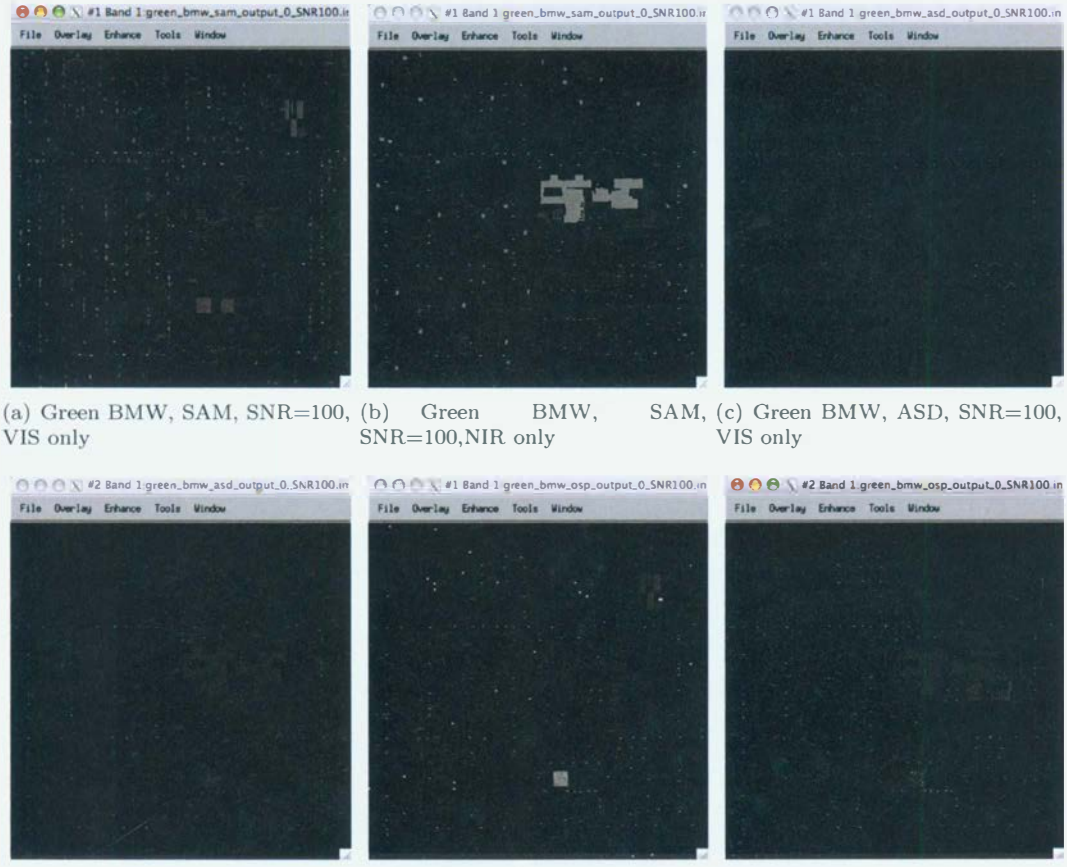
(a) Blue Ford Focus, CEM, SNR=100, VIS only

(b) Blue Ford Focus, CEM, SNR=100, NIR only

(c) Blue Ford Focus, ACE, SNR=100, VIS only

(d) Blue Ford Focus, ACE, SNR=100, NIR only

Figure 7.14: detection images for the scene with blue Ford Focuses parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the output of CEM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ACE using only the VIS and NIR channels respectively. The histogram of these output images were all stretched to display roughly the highest 5% of detections.



(a) Green BMW, SAM, SNR=100, VIS only

(b) Green BMW, SAM, SNR=100,NIR only

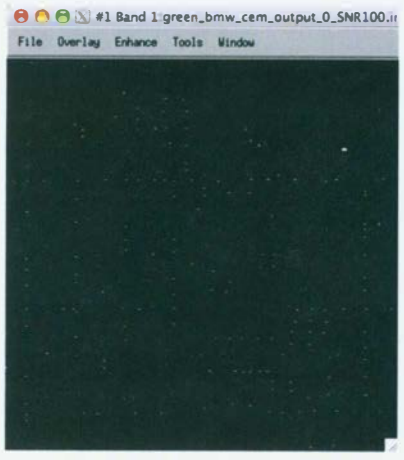
(c) Green BMW, ASD, SNR=100, VIS only

(d) Green BMW, ASD, SNR=100, NIR only

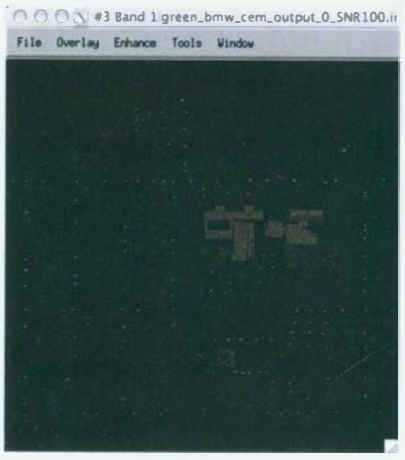
(e) Green BMW, OSP, SNR=100, VIS only

(f) Green BMW, OSP, SNR=100, NIR only

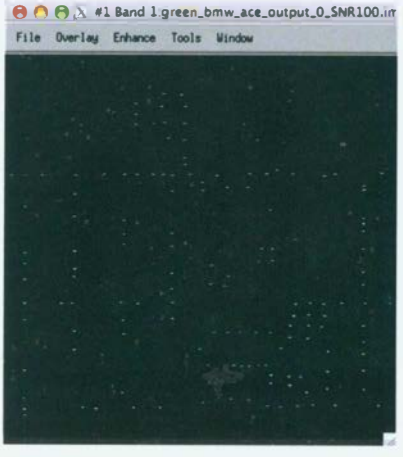
Figure 7.15: detection images for the scene with green BMWs parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the inverted output of SAM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ASD using only the VIS and NIR channels respectively and Figures (e) and (f) show the output of OSP in the same manner. The histogram of these output images were all stretched to display roughly the highest 5% of detections.



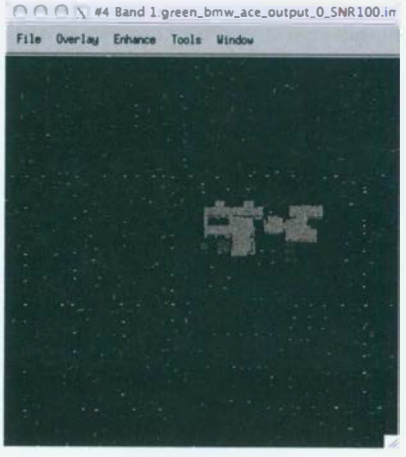
(a) Green BMW, CEM, SNR=100, VIS only



(b) Green BMW, CEM, SNR=100, NIR only



(c) Green BMW, ACE, SNR=100, VIS only



(d) Green BMW, ACE, SNR=100, NIR only

Figure 7.16: detection images for the scene with green BMWs parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the output of CEM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ACE using only the VIS and NIR channels respectively. The histogram of these output images were all stretched to display roughly the highest 5% of detections.

(a) White Saturn Vue, SAM, SNR=100, VIS only

(b) White Saturn Vue, SAM, SNR=100,NIR only

(c) White Saturn Vue, ASD, SNR=100, VIS only

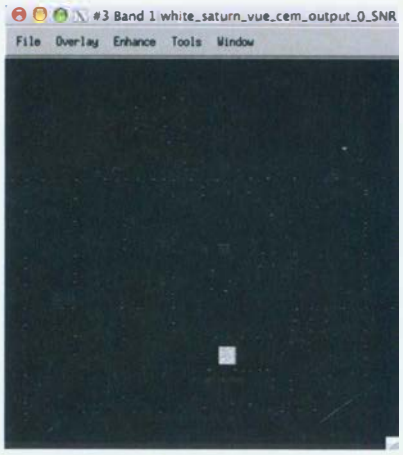
(d) White Saturn Vue, ASD, SNR=100, NIR only

(e) White Saturn Vue, OSP, SNR=100, VIS only

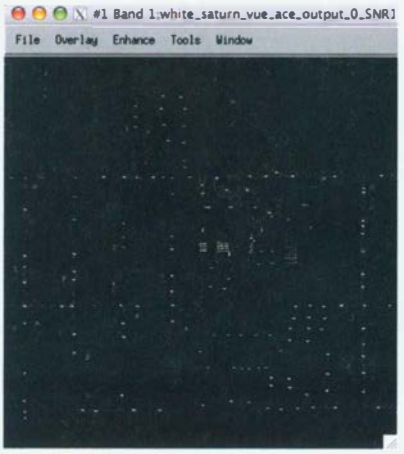
(f) White Saturn Vue, OSP, SNR=100, NIR only

Figure 7.17: detection images for the scene with White Saturn Vues parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the inverted output of SAM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ASD using only the VIS and NIR channels respectively and Figures (e) and (f) show the output of OSP in the same manner.. The histogram of these output images were all stretched to display roughly the highest 5% of detections.

(a) White Saturn Vue, CEM, SNR=100, VIS only



(b) White Saturn Vue, CEM, SNR=100, NIR only



(c) White Saturn Vue, ACE, SNR=100, VIS only

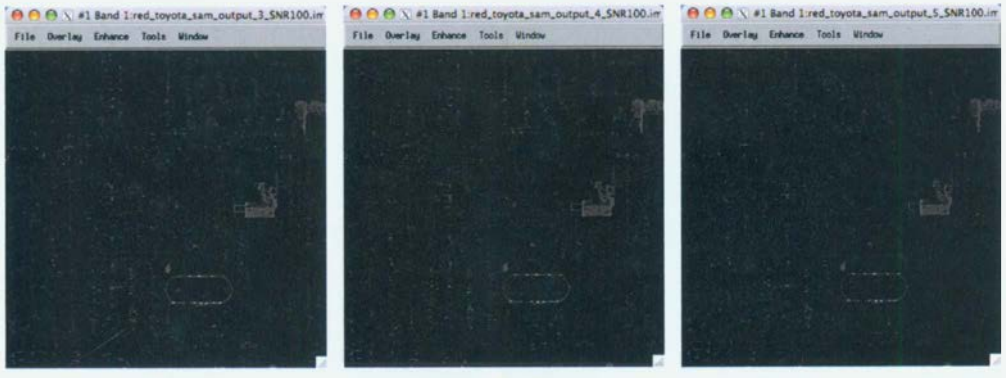
(d) White Saturn Vue, ACE, SNR=100, NIR only

Figure 7.18: detection images for the scene with White Saturn Vues parked throughout. The Visible and Near-infrared channels were used separately to create these detection images. Images (a) and (b) show the output of CEM using only the VIS and NIR channels respectively. Figures (c) and (d) show the output of ACE using only the VIS and NIR channels respectively. The histogram of these output images were all stretched to display roughly the highest 5% of detections.

(a) Red Toyota, SAM, SNR=100, 0 pixel shift

(b) Red Toyota, SAM, SNR=100, 0.1 pixel shift

(c) Red Toyota, SAM, SNR=100, 0.2 pixel shift

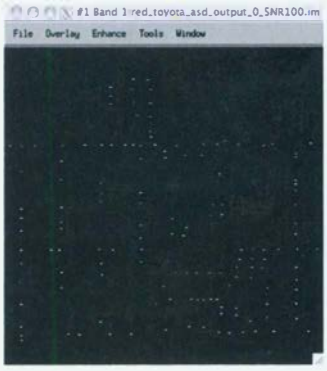


(d) Red Toyota, SAM, SNR=100, 0.3 pixel shift

(e) Red Toyota, SAM, SNR=100, 0.4 pixel shift

(f) Red Toyota, SAM, SNR=100, 0.5 pixel shift

Figure 7.19: Results from running SAM on the scene with red Toyotas as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Red Toyota, ASD, SNR=100, 0 pixel shift

(b) Red Toyota, ASD, SNR=100, 0.1 pixel shift

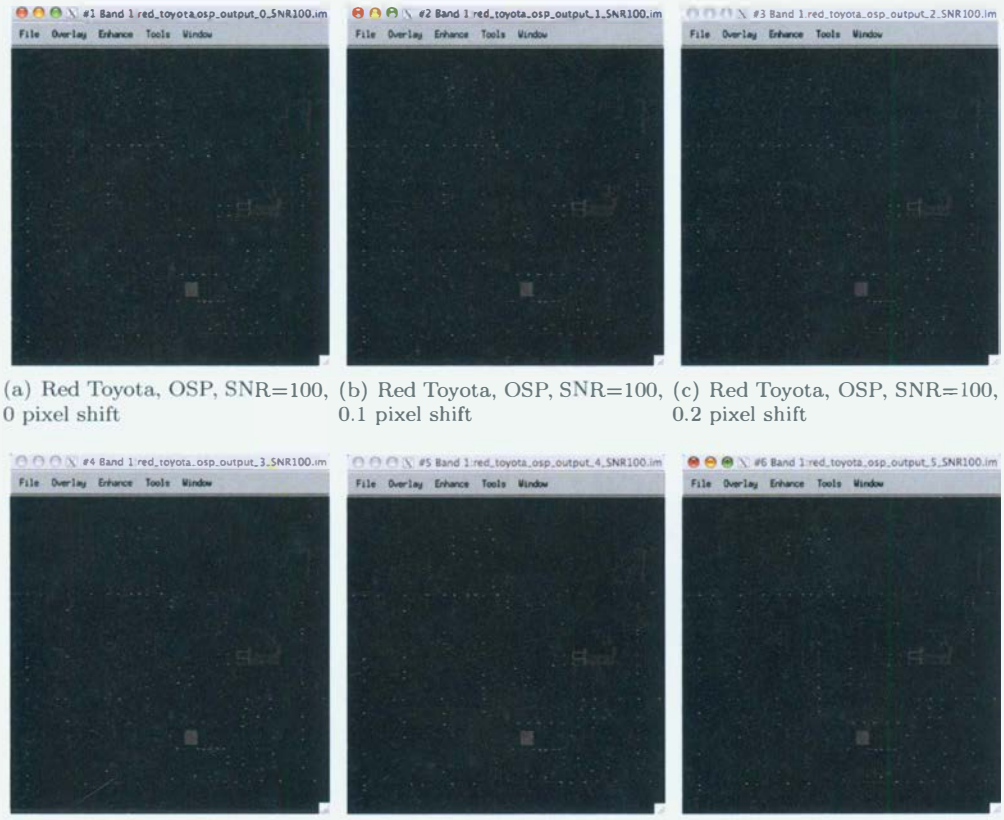
(c) Red Toyota, ASD, SNR=100, 0.2 pixel shift

(d) Red Toyota, ASD, SNR=100, 0.3 pixel shift

(e) Red Toyota, ASD, SNR=100, 0.4 pixel shift

(f) Red Toyota, ASD, SNR=100, 0.5 pixel shift

Figure 7.20: Results from running ASD on the scene with red Toyotas as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Red Toyota, OSP, SNR=100, 0 pixel shift

(b) Red Toyota, OSP, SNR=100, 0.1 pixel shift

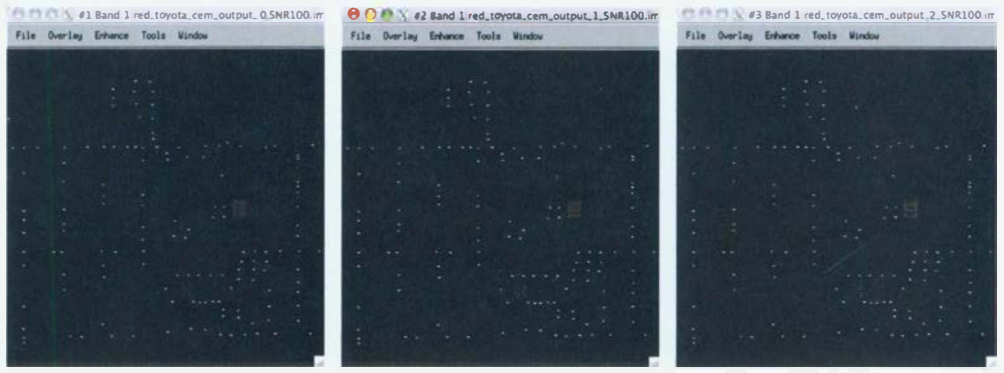
(c) Red Toyota, OSP, SNR=100, 0.2 pixel shift

(d) Red Toyota, OSP, SNR=100, 0.3 pixel shift

(e) Red Toyota, OSP, SNR=100, 0.4 pixel shift

(f) Red Toyota, OSP, SNR=100, 0.5 pixel shift

Figure 7.21: Results from running OSP on the scene with red Toyotas as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Red Toyota, CEM, SNR=100, 0 pixel shift

(b) Red Toyota, CEM, SNR=100, 0.1 pixel shift

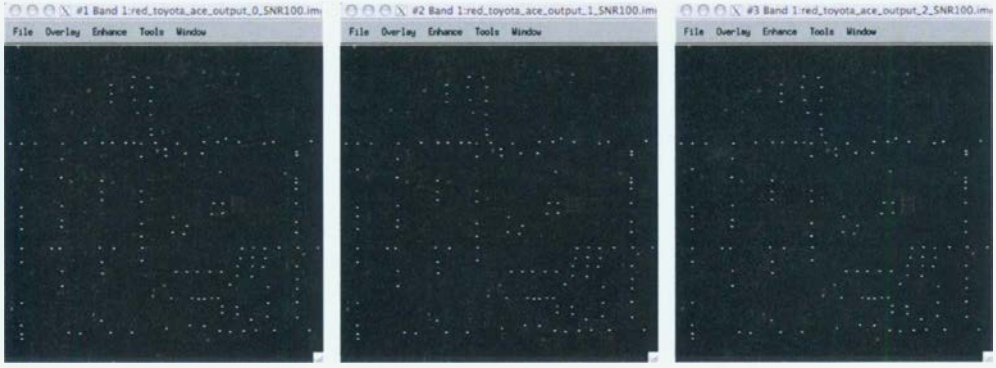
(c) Red Toyota, CEM, SNR=100, 0.2 pixel shift

(d) Red Toyota, CEM, SNR=100, 0.3 pixel shift

(e) Red Toyota, CEM, SNR=100, 0.4 pixel shift

(f) Red Toyota, CEM, SNR=100, 0.5 pixel shift

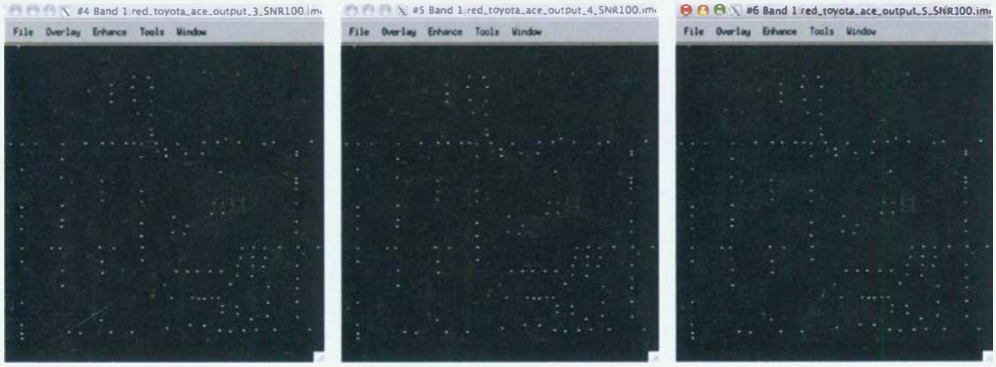
Figure 7.22: Results from running CEM on the scene with red Toyotas as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Red Toyota, ACE, SNR=100, 0 pixel shift

(b) Red Toyota, ACE, SNR=100, 0.1 pixel shift

(c) Red Toyota, ACE, SNR=100, 0.2 pixel shift

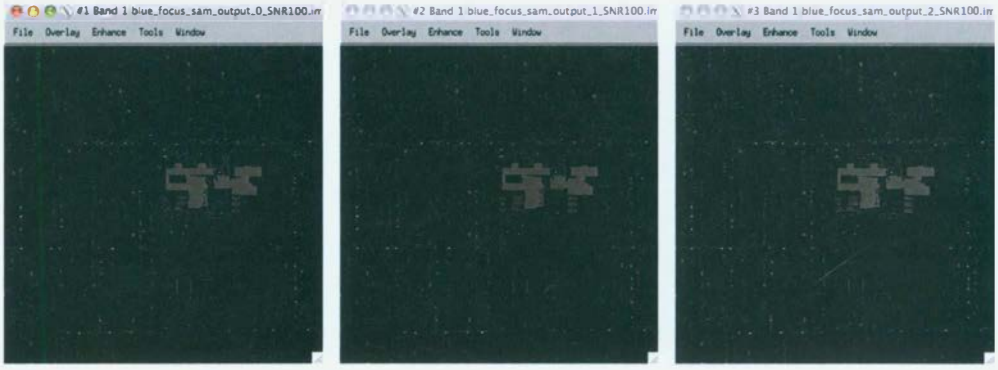


(d) Red Toyota, ACE, SNR=100, 0.3 pixel shift

(e) Red Toyota, ACE, SNR=100, 0.4 pixel shift

(f) Red Toyota, ACE, SNR=100, 0.5 pixel shift

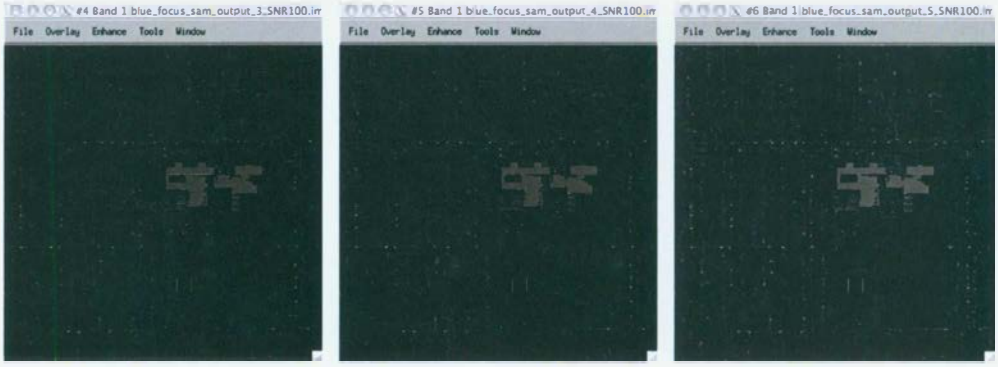
Figure 7.23: Results from running ACE on the scene with red Toyotas as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Blue Focus, SAM, SNR=100, 0 pixel shift

(b) Blue Focus, SAM, SNR=100, 0.1 pixel shift

(c) Blue Focus, SAM, SNR=100, 0.2 pixel shift

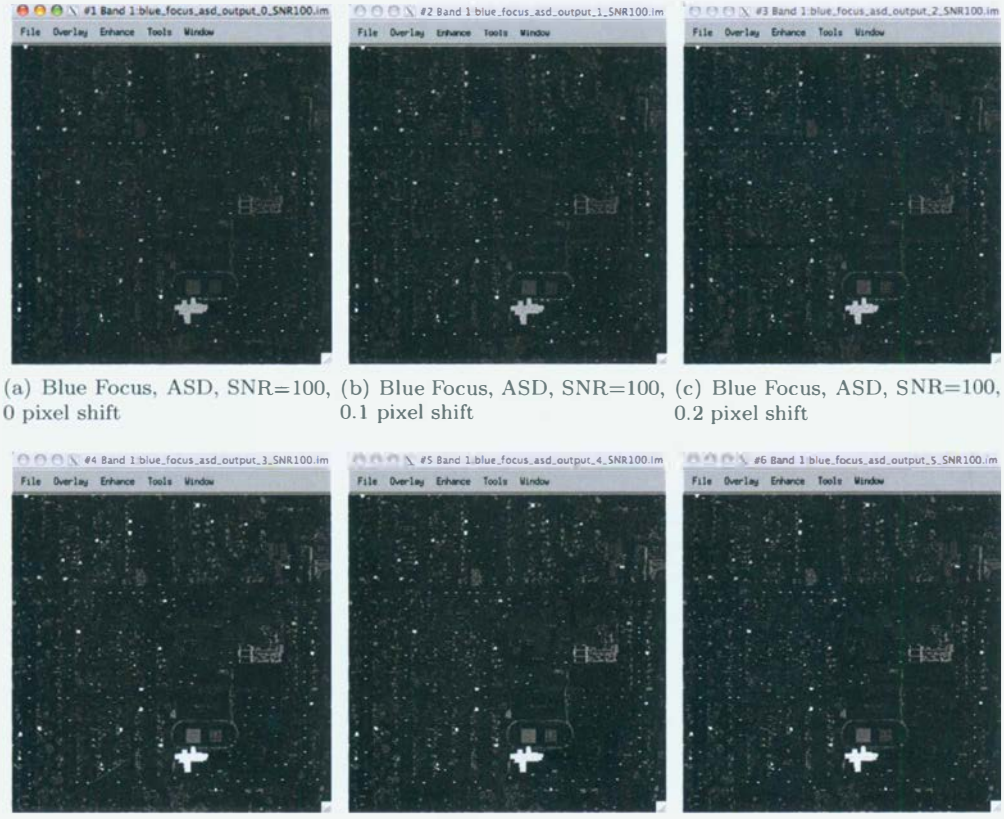


(d) Blue Focus, SAM, SNR=100, 0.3 pixel shift

(e) Blue Focus, SAM, SNR=100, 0.4 pixel shift

(f) Blue Focus, SAM, SNR=100, 0.5 pixel shift

Figure 7.24: Results from running SAM on the scene with blue Ford Focuses as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Blue Focus, ASD, SNR=100, 0 pixel shift

(b) Blue Focus, ASD, SNR=100, 0.1 pixel shift

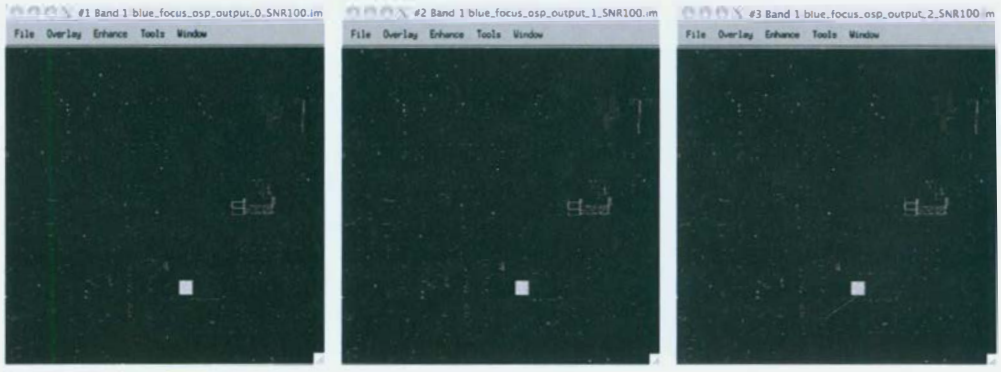
(c) Blue Focus, ASD, SNR=100, 0.2 pixel shift

(d) Blue Focus, ASD, SNR=100, 0.3 pixel shift

(e) Blue Focus, ASD, SNR=100, 0.4 pixel shift

(f) Blue Focus, ASD, SNR=100, 0.5 pixel shift

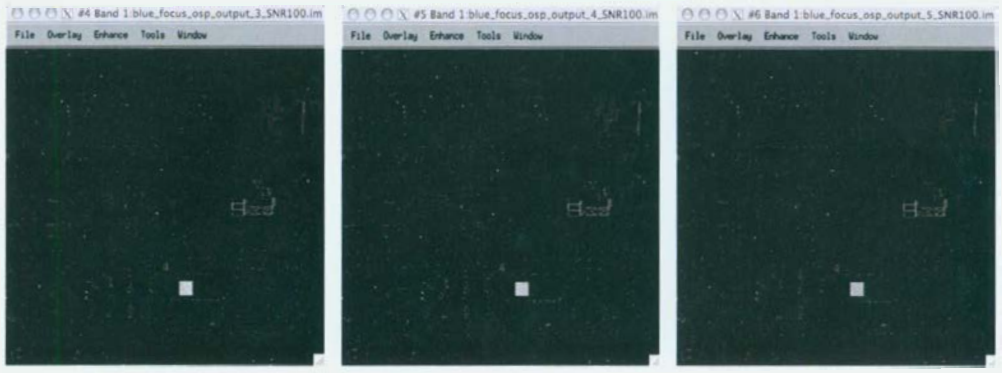
Figure 7.25: Results from running ASD on the scene with blue Ford Focuses as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Blue Focus, OSP, SNR=100, 0 pixel shift

(b) Blue Focus, OSP, SNR=100, 0.1 pixel shift

(c) Blue Focus, OSP, SNR=100, 0.2 pixel shift

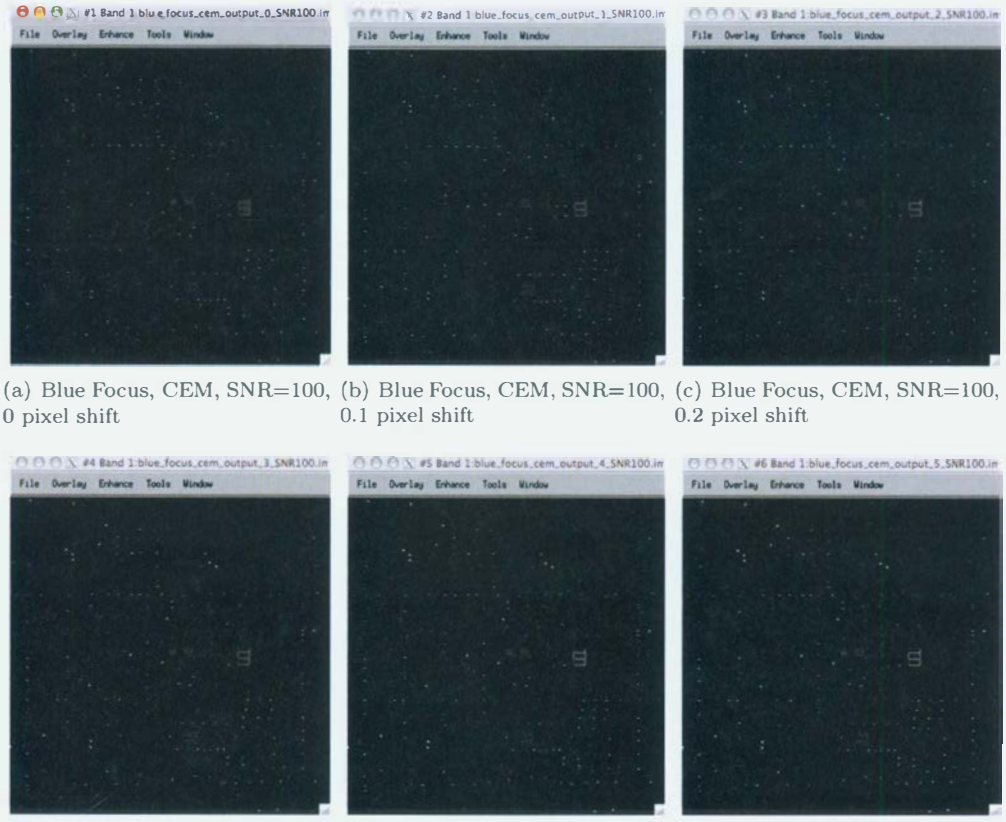


(d) Blue Focus, OSP, SNR=100, 0.3 pixel shift

(e) Blue Focus, OSP, SNR=100, 0.4 pixel shift

(f) Blue Focus, OSP, SNR=100, 0.5 pixel shift

Figure 7.26: Results from running OSP on the scene with blue Ford Focuses as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Blue Focus, CEM, SNR=100, 0 pixel shift

(b) Blue Focus, CEM, SNR=100, 0.1 pixel shift

(c) Blue Focus, CEM, SNR=100, 0.2 pixel shift

(d) Blue Focus, CEM, SNR=100, 0.3 pixel shift

(e) Blue Focus, CEM, SNR=100, 0.4 pixel shift

(f) Blue Focus, CEM, SNR=100, 0.5 pixel shift

Figure 7.27: Results from running CEM on the scene with blue Ford Focuses as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

(a) Blue Focus, ACE, SNR=100, 0 pixel shift

(b) Blue Focus, ACE, SNR=100, 0.1 pixel shift

(c) Blue Focus, ACE, SNR=100, 0.2 pixel shift

(d) Blue Focus, ACE, SNR=100, 0.3 pixel shift

(e) Blue Focus, ACE, SNR=100, 0.4 pixel shift

(f) Blue Focus, ACE, SNR=100, 0.5 pixel shift

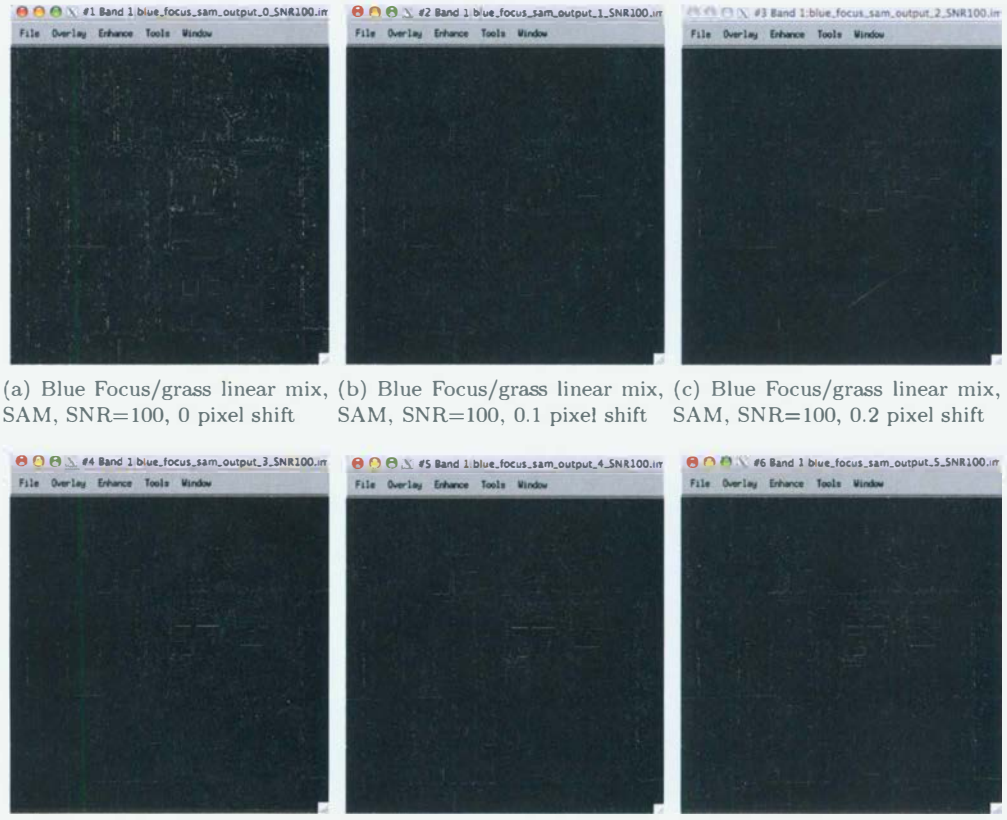
Figure 7.28: Results from running ACE on the scene with blue Ford Focuses as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

Looking at some of the detection images for the blue Ford Focus, it is quite noticeable that vehicles parked in the grass area are not as easily detectable as vehicles parked on road surfaces, especially as misregistration increases. In an attempt to improve the detectability of vehicles parked in a grass background, a linear mixing model was used in only the NIR set of bands to predict the target signature in misregistered imagery. The spectral mixtures consisted of only the grass field and the blue Ford Focus in only the NIR bands. This mixture model that was used can be seen in equation 7.2

$$d_{VIS} = focus_{VIS} \tag{7.1}$$

$$d_{NIR} = \alpha focus_{NIR} + (1-\alpha) grass_{NIR} \tag{7.2}$$

where $focus_{VIS}$ is the reflectance of the blue Ford Focus in the VIS, $focus_{NIR}$ is the reflectance of the blue Ford Focus in the NIR, and $grass_{NIR}$ is the reflectance of the grass in the NIR. Also, here $\alpha = 0.25$, as we've assumed the amount of misregistration between the VIS and NIR bands is between 0 and 0.5, so 0.25 is used here to provide a best guess estimate of the misregistration amount. A few results for the blue Ford Focus using the linear mixture model can be seen in Figures 7.29, 7.30, and 7.31. It can be seen that the vehicles in the detection maps stand out a bit more in the grass field for larger amounts of misregistration, however, more false alarms appear throughout the detection map, and vehicles parked against the other background surfaces are not detected as well. ROC curves for these detection maps can be seen in Figures 7.52, 7.53 and 7.54. Similar results were produced for the other algorithms, OSP and CEM but are not shown here.



(a) Blue Focus/grass linear mix, SAM, SNR=100, 0 pixel shift

(b) Blue Focus/grass linear mix, SAM, SNR=100, 0.1 pixel shift

(c) Blue Focus/grass linear mix, SAM, SNR=100, 0.2 pixel shift

(d) Blue Focus/grass linear mix, SAM, SNR=100, 0.3 pixel shift

(e) Blue Focus/grass linear mix, SAM, SNR=100, 0.4 pixel shift

(f) Blue Focus/grass linear mix, SAM, SNR=100, 0.5 pixel shift

Figure 7.29: Results from running SAM on the scene with blue Ford Focuses as targets at different amounts of misregistration. A linear mixture of vehicle and grass was used in the NIR spectral bands to attempt to compensate for band-to-band misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

(a) Blue Focus/grass linear mix, ASD, SNR=100, 0 pixel shift

(b) Blue Focus/grass linear mix, ASD, SNR=100, 0.1 pixel shift

(c) Blue Focus/grass linear mix, ASD, SNR=100, 0.2 pixel shift

(d) Blue Focus/grass linear mix, ASD, SNR=100, 0.3 pixel shift

(e) Blue Focus/grass linear mix, ASD, SNR=100, 0.4 pixel shift

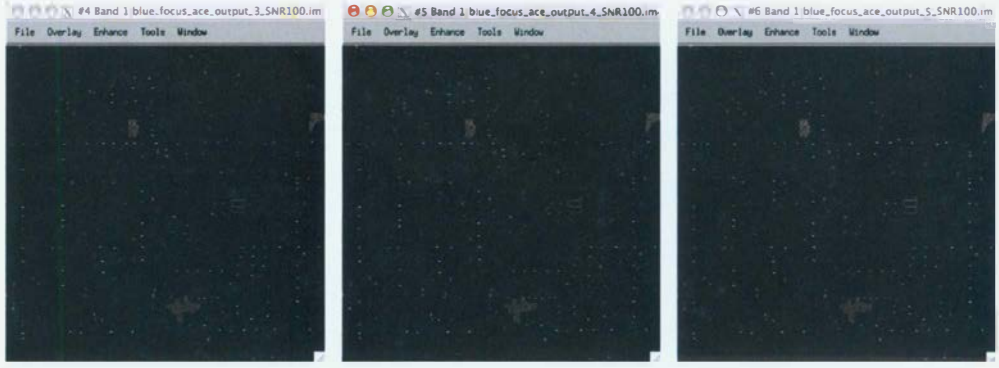
(f) Blue Focus/grass linear mix, ASD, SNR=100, 0.5 pixel shift

Figure 7.30: Results from running ASD on the scene with blue Ford Focuses as targets at different amounts of misregistration. A linear mixture of vehicle and grass was used in the NIR spectral bands to attempt to compensate for band-to-band misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

(a) Blue Focus/grass linear mix, ACE, SNR=100, 0 pixel shift

(b) Blue Focus/grass linear mix, ACE, SNR=100, 0.1 pixel shift

(c) Blue Focus/grass linear mix, ACE, SNR=100, 0.2 pixel shift

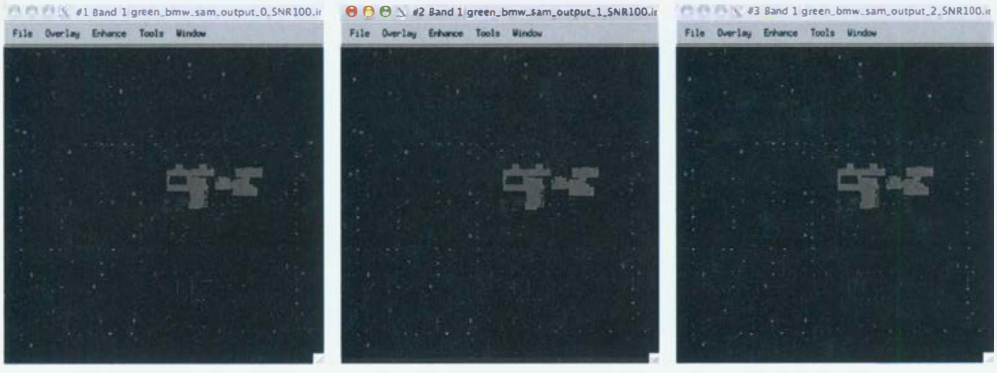


(d) Blue Focus/grass linear mix, ACE, SNR=100, 0.3 pixel shift

(e) Blue Focus/grass linear mix, ACE, SNR=100, 0.4 pixel shift

(f) Blue Focus/grass linear mix, ACE, SNR=100, 0.5 pixel shift

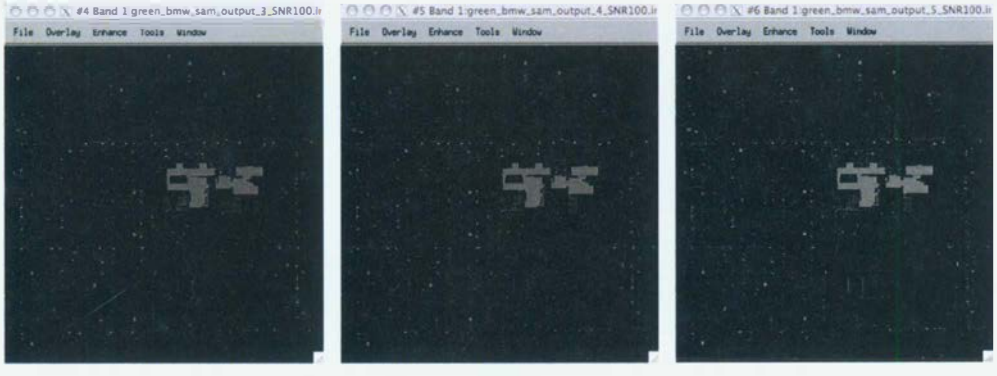
Figure 7.31: Results from running ACE on the scene with blue Ford Focuses as targets at different amounts of misregistration. A linear mixture of vehicle and grass was used in the NIR spectral bands to attempt to compensate for band-to-band misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Green BMW, SAM, SNR=100, 0 pixel shift

(b) Green BMW, SAM, SNR=100, 0.1 pixel shift

(c) Green BMW, SAM, SNR=100, 0.2 pixel shift

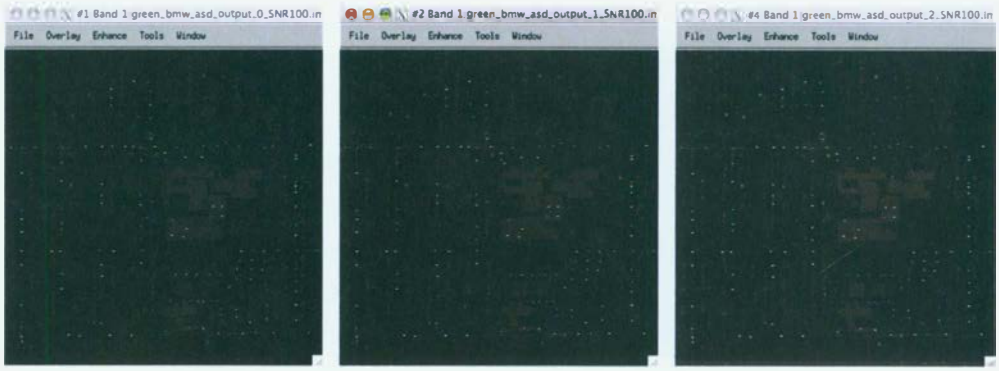


(d) Green BMW, SAM, SNR=100, 0.3 pixel shift

(e) Green BMW, SAM, SNR=100, 0.4 pixel shift

(f) Green BMW, SAM, SNR=100, 0.5 pixel shift

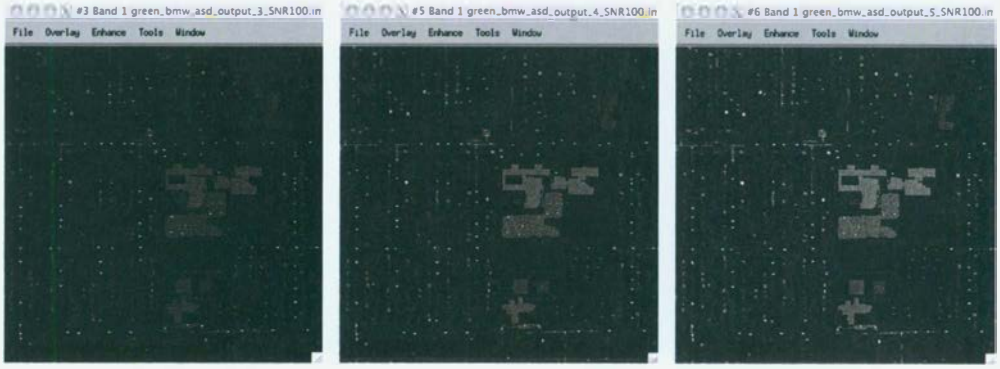
Figure 7.32: Results from running SAM on the scene with green BMWs as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Green BMW, ASD, SNR=100, 0 pixel shift

(b) Green BMW, ASD, SNR=100, 0.1 pixel shift

(c) Green BMW, ASD, SNR=100, 0.2 pixel shift

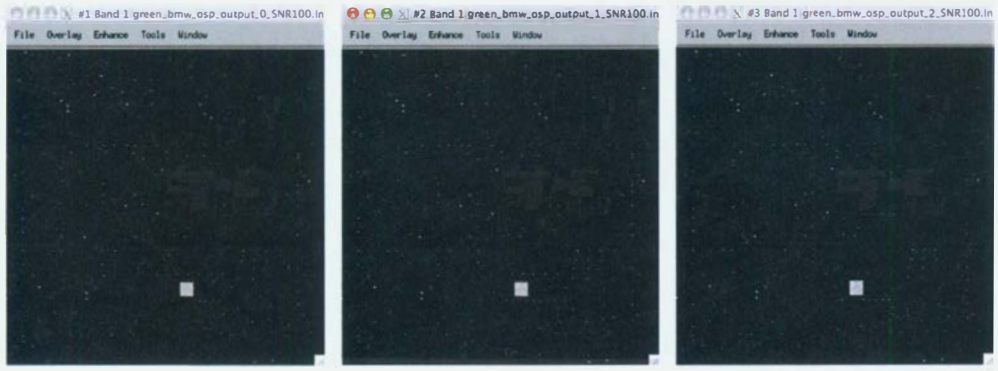


(d) Green BMW, ASD, SNR=100, 0.3 pixel shift

(e) Green BMW, ASD, SNR=100, 0.4 pixel shift

(f) Green BMW, ASD, SNR=100, 0.5 pixel shift

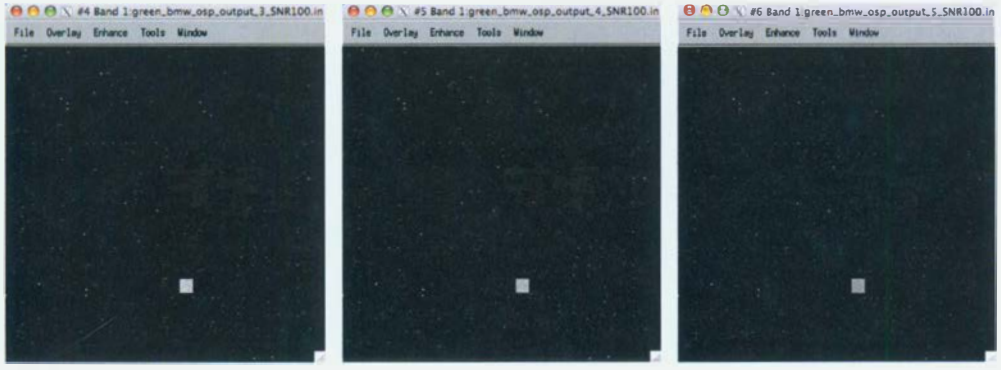
Figure 7.33: Results from running ASD on the scene with green BMWs as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Green BMW, OSP, SNR=100, 0 pixel shift

(b) Green BMW, OSP, SNR=100, 0.1 pixel shift

(c) Green BMW, OSP, SNR=100, 0.2 pixel shift

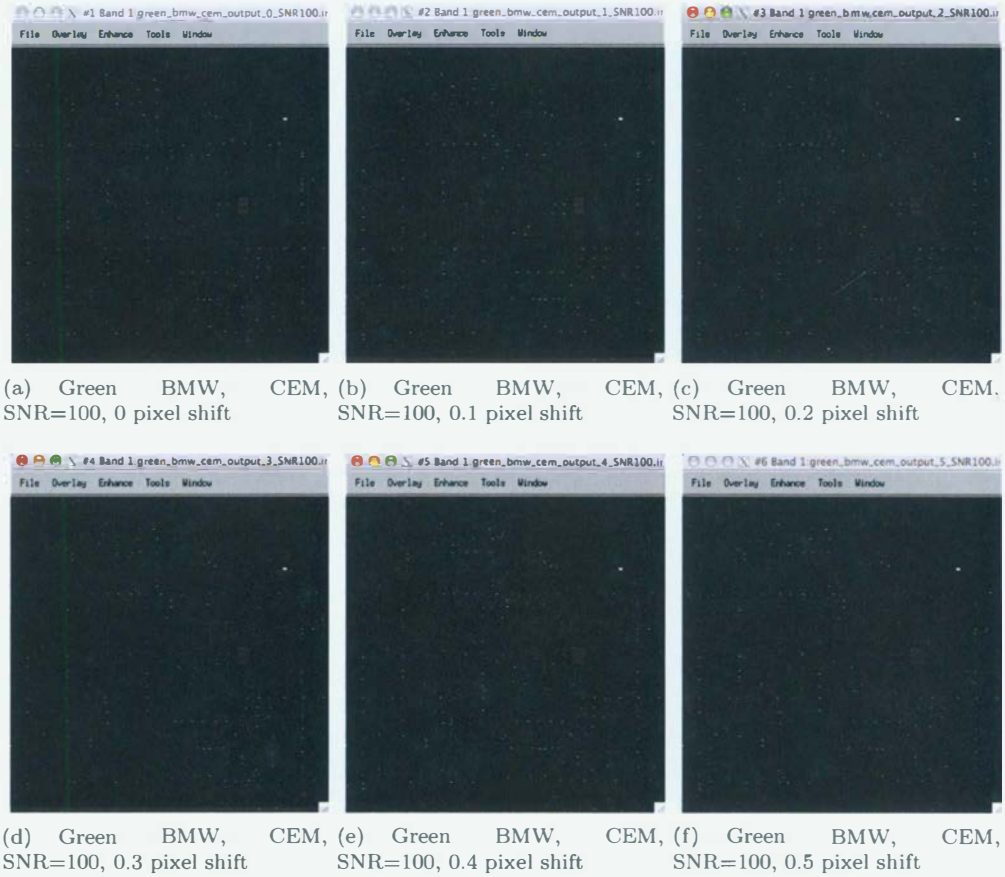


(d) Green BMW, OSP, SNR=100, 0.3 pixel shift

(e) Green BMW, OSP, SNR=100, 0.4 pixel shift

(f) Green BMW, OSP, SNR=100, 0.5 pixel shift

Figure 7.34: Results from running OSP on the scene with green BMWs as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Green BMW, CEM, SNR=100, 0 pixel shift

(b) Green BMW, CEM, SNR=100, 0.1 pixel shift

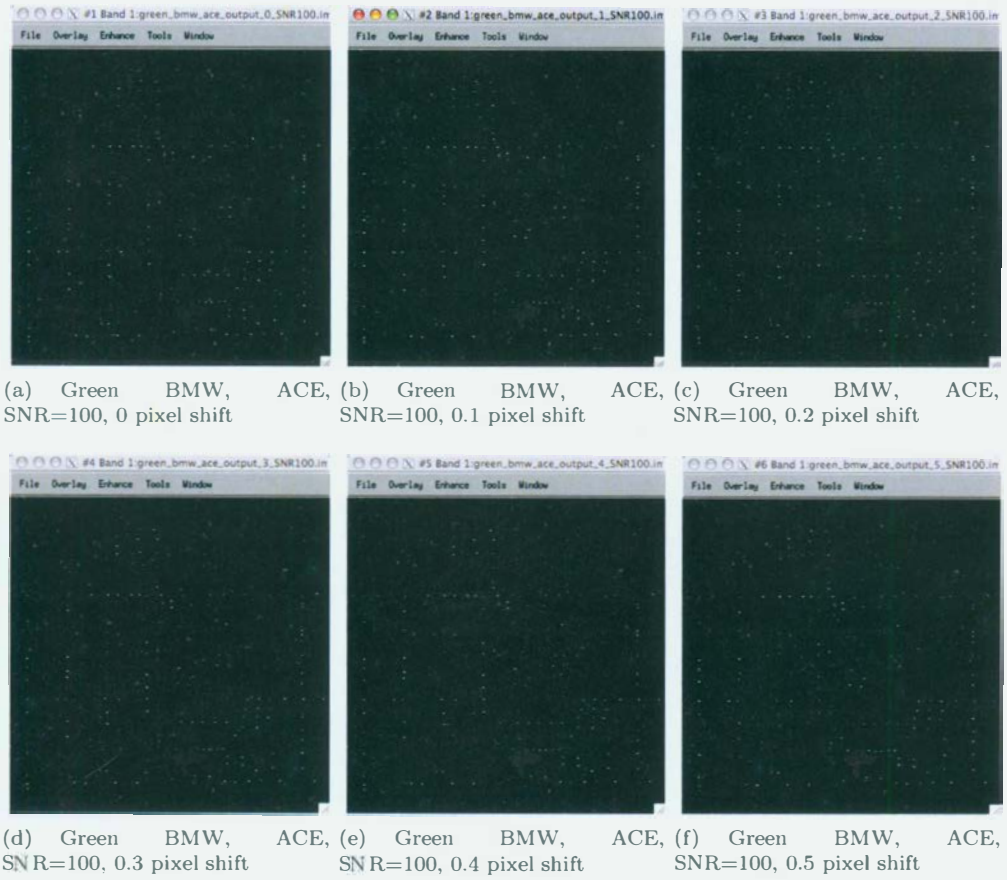
(c) Green BMW, CEM, SNR=100, 0.2 pixel shift

(d) Green BMW, CEM, SNR=100, 0.3 pixel shift

(e) Green BMW, CEM, SNR=100, 0.4 pixel shift

(f) Green BMW, CEM, SNR=100, 0.5 pixel shift

Figure 7.35: Results from running CEM on the scene with green BMWs as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) Green BMW, ACE, SNR=100, 0 pixel shift

(b) Green BMW, ACE, SNR=100, 0.1 pixel shift

(c) Green BMW, ACE, SNR=100, 0.2 pixel shift

(d) Green BMW, ACE, SNR=100, 0.3 pixel shift

(e) Green BMW, ACE, SNR=100, 0.4 pixel shift

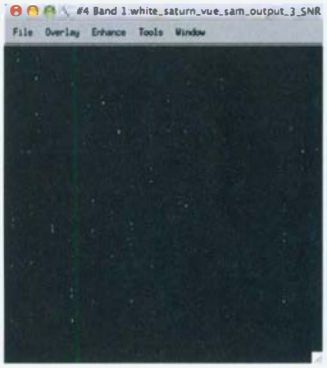
(f) Green BMW, ACE, SNR=100, 0.5 pixel shift

Figure 7.36: Results from running ACE on the scene with green BMWs as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

(a) White Saturn Vue, SAM, SNR=100, 0 pixel shift

(b) White Saturn Vue, SAM, SNR=100, 0.1 pixel shift

(c) White Saturn Vue, SAM, SNR=100, 0.2 pixel shift

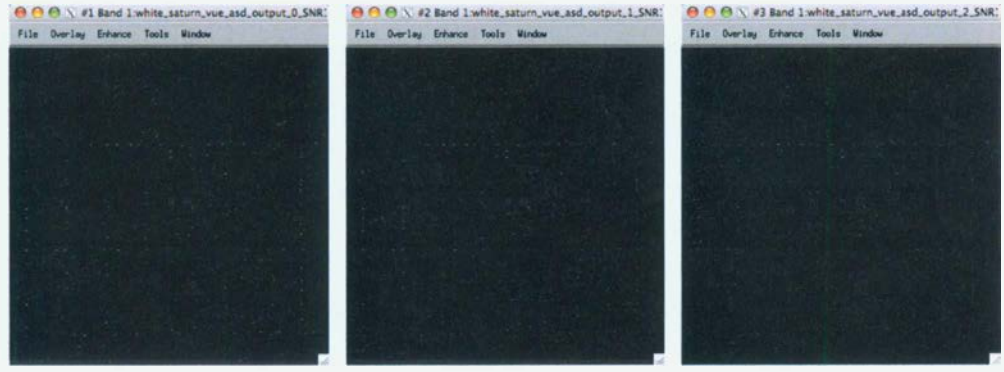


(d) White Saturn Vue, SAM, SNR=100, 0.3 pixel shift

(e) White Saturn Vue, SAM, SNR=100, 0.4 pixel shift

(f) White Saturn Vue, SAM, SNR=100, 0.5 pixel shift

Figure 7.37: Results from running SAM on the scene with White Saturn Vues as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) White Saturn Vue, ASD, SNR=100, 0 pixel shift

(b) White Saturn Vue, ASD, SNR=100, 0.1 pixel shift

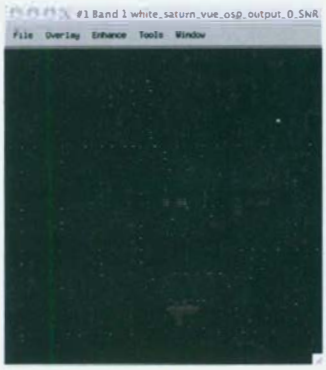
(c) White Saturn Vue, ASD, SNR=100, 0.2 pixel shift

(d) White Saturn Vue, ASD, SNR=100, 0.3 pixel shift

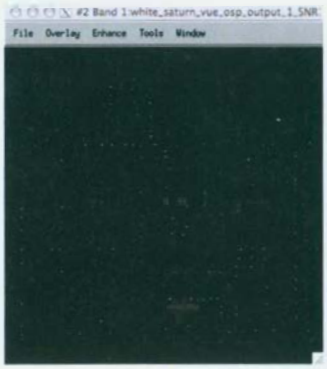
(e) White Saturn Vue, ASD, SNR=100, 0.4 pixel shift

(f) White Saturn Vue, ASD, SNR=100, 0.5 pixel shift

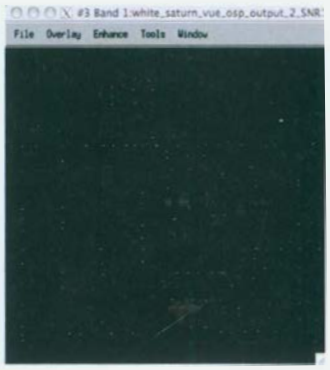
Figure 7.38: Results from running ASD on the scene with White Saturn Vues as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) White Saturn Vue, OSP, SNR=100, 0 pixel shift



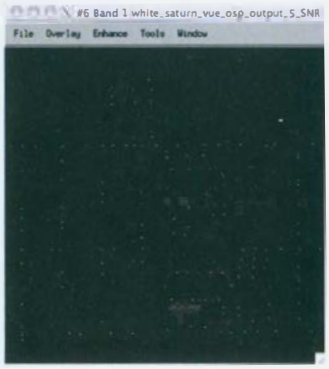
(b) White Saturn Vue, OSP, SNR=100, 0.1 pixel shift



(c) White Saturn Vue, OSP, SNR=100, 0.2 pixel shift

(d) White Saturn Vue, OSP, SNR=100, 0.3 pixel shift

(e) White Saturn Vue, OSP, SNR=100, 0.4 pixel shift



(f) White Saturn Vue, OSP, SNR=100, 0.5 pixel shift

Figure 7.39: Results from running OSP on the scene with White Saturn Vues as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

(a) White Saturn Vue, CEM, SNR=100, 0 pixel shift

(b) White Saturn Vue, CEM, SNR=100, 0.1 pixel shift

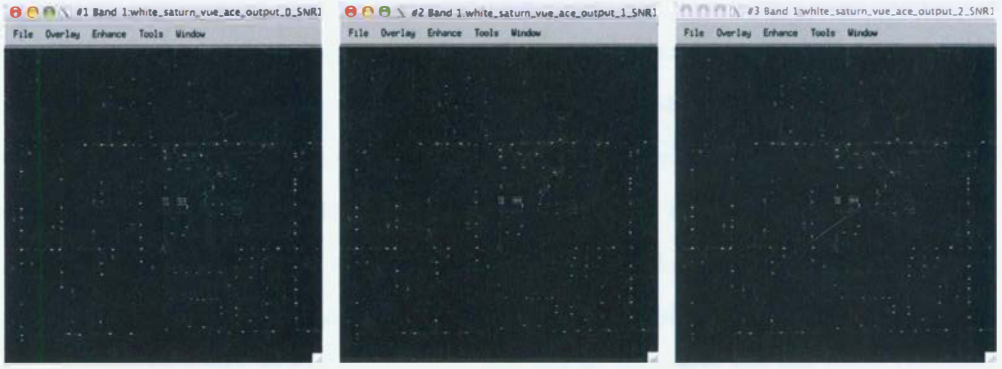
(c) White Saturn Vue, CEM, SNR=100, 0.2 pixel shift

(d) White Saturn Vue, CEM, SNR=100, 0.3 pixel shift

(e) White Saturn Vue, CEM, SNR=100, 0.4 pixel shift

(f) White Saturn Vue, CEM, SNR=100, 0.5 pixel shift

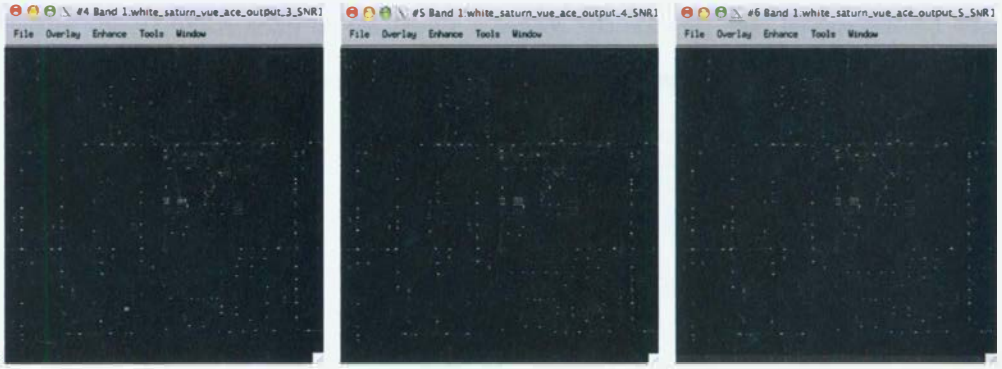
Figure 7.40: Results from running CEM on the scene with White Saturn Vues as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.



(a) White Saturn Vue, ACE, SNR=100, 0 pixel shift

(b) White Saturn Vue, ACE, SNR=100, 0.1 pixel shift

(c) White Saturn Vue, ACE, SNR=100, 0.2 pixel shift



(d) White Saturn Vue, ACE, SNR=100, 0.3 pixel shift

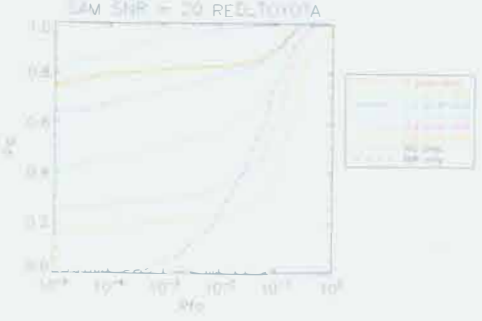
(e) White Saturn Vue, ACE, SNR=100, 0.4 pixel shift

(f) White Saturn Vue, ACE, SNR=100, 0.5 pixel shift

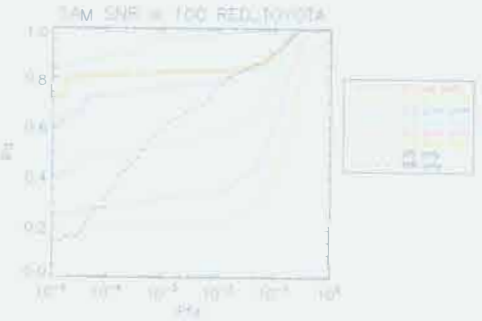
Figure 7.41: Results from running ACE on the scene with White Saturn Vues as targets at different amounts of misregistration. The amount of misregistration increases in increments of 0.1 pixel, from no misregistration in image (a) and 0.5 pixel misregistration in image(f). All images show roughly the top 5% of pixels.

After detection maps were generated using the SAM, ASD, OSP, CEM and ACE algorithms on all the synthetic scenes, ROC curves were generated to show the probability of detection at a given false alarm rate. The first step in generating the ROC curves was looking for target pixels in the degraded imagery. A true color image was loaded, target pixel locations were recorded. A mask was also created around pixels nearby the target that could not be classified as either target or non-target. The detection values were recorded for each target pixel, and the number of non-target pixels with a detection value greater than a target pixel were counted. Using these numbers, the probability of detection for every false alarm rate was computed, and could be plotted to generate a ROC curve. These ROC curves can be seen in Figures 7.42 through 7.64. Plotted in each of these figures are the probability of detection at a given false alarm rate at each amount of misregistration, from 0 to 0.5 pixel, as well as the detection rates using only the VIS and NIR bands. This provides an easy way to compare algorithms, as well as compare the performance using the full spectrum vs the performance using VIS and NIR bands separately. Three sets of ROC curves for the three images with different SNRs are plotted for each figure.

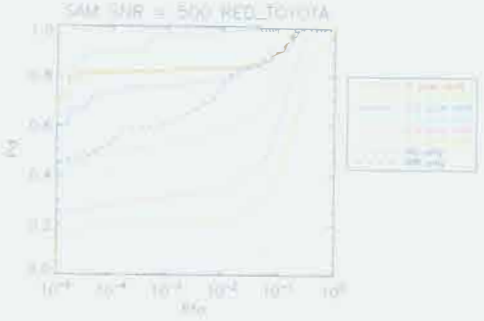
The results shown here provide the reader a visual aid to understand trends in detection results. All these results are interpreted and discussed in Chapter 8



(a) Red Toyota, SAM, SNR=20

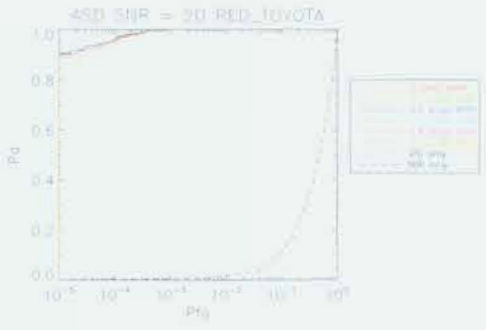


(b) Red Toyota, SAM, SNR=100

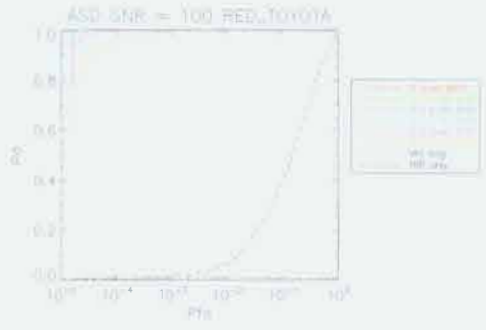


(c) Red Toyota, SAM, SNR=500

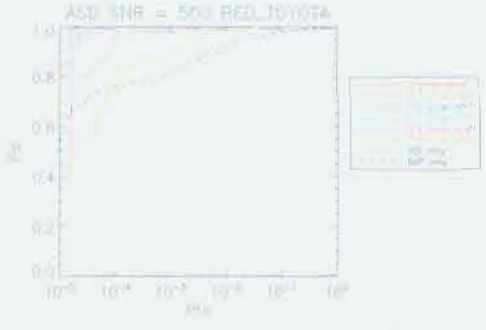
Figure 7.42: ROC curves for red Toyota using SAM.



(a) Red Toyota, ASD, SNR=20

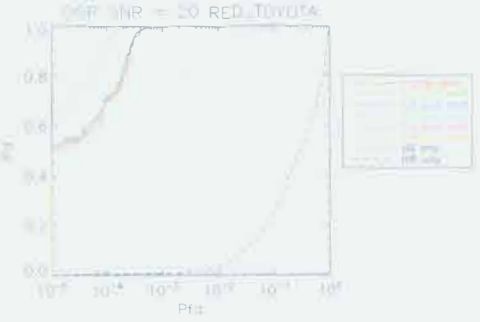


(b) Red Toyota, ASD, SNR=100

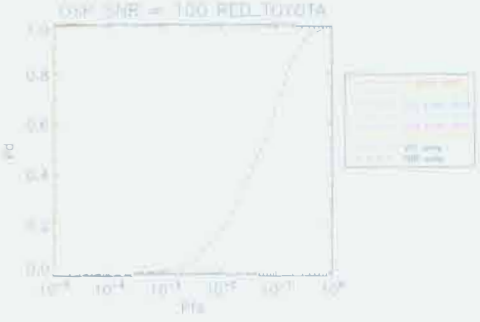


(c) Red Toyota, ASD, SNR=500

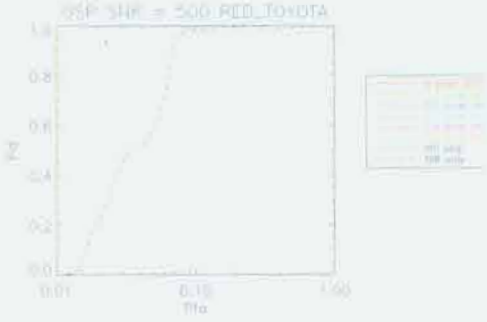
Figure 7.43: R●C curves for red Toyota using ASD.



(a) Red Toyota, OSP, SNR=20

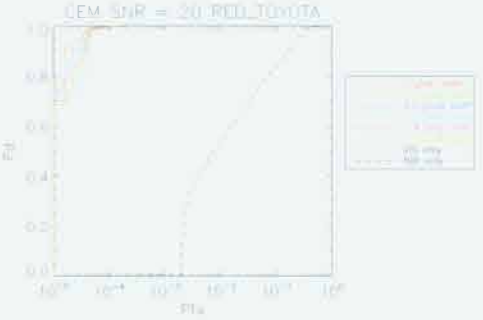


(b) Red Toyota, OSP, SNR=100

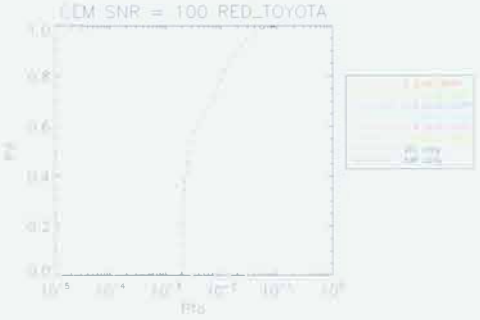


(c) Red Toyota, OSP, SNR=500

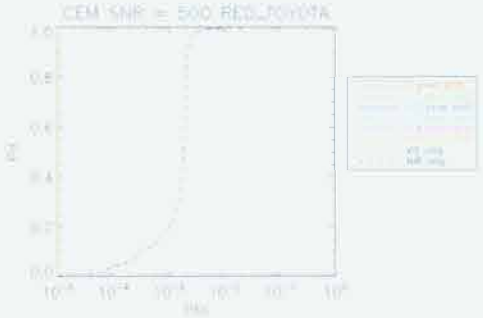
Figure 7.44: ROC curves for red Toyota using OSP.



(a) Red Toyota, CEM, SNR=20

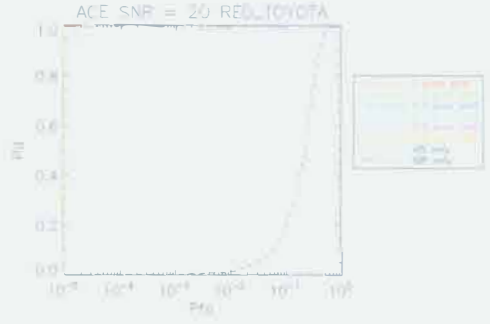


(b) Red Toyota, CEM, SNR=100

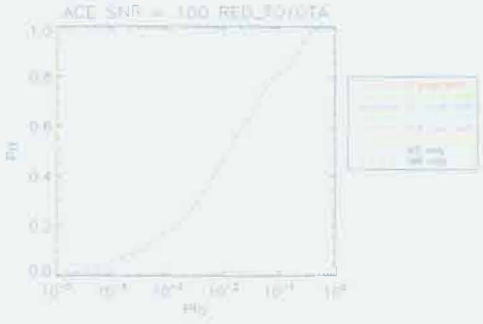


(c) Red Toyota, CEM, SNR=500

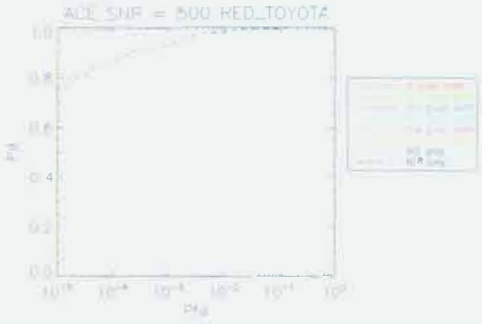
Figure 7.45: ROC curves for red Toyota using CEM.



(a) Red Toyota, ACE, SNR=20

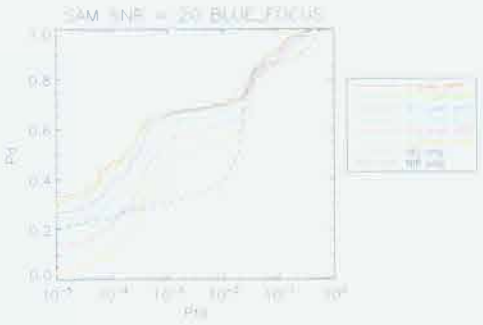


(b) Red Toyota, ACE, SNR=100

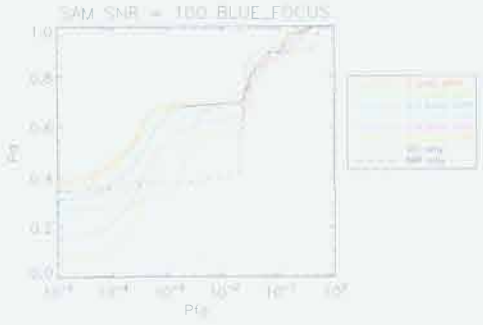


(c) Red Toyota, ACE, SNR=500

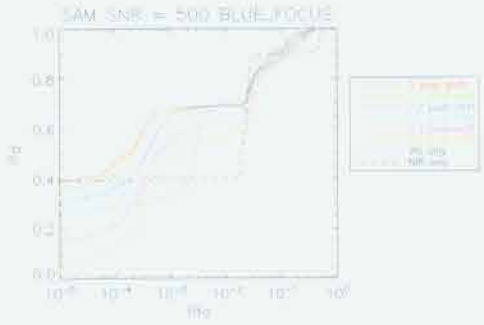
Figure 7.46: ROC curves for red Toyota using ACE.



(a) Blue Focus, SAM, SNR=20

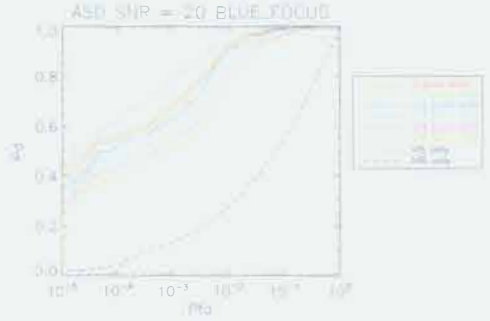


(b) Blue Focus, SAM, SNR=100

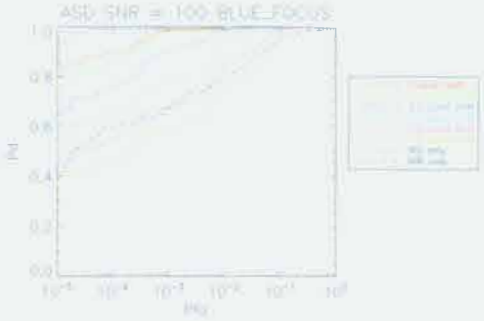


(c) Blue Focus, SAM, SNR=500

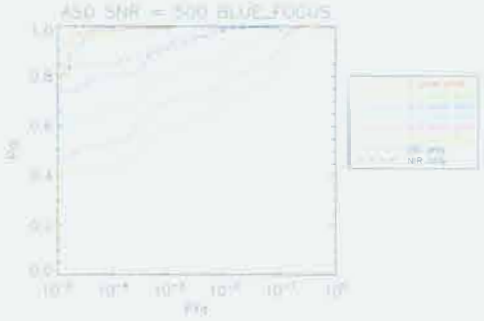
Figure 7.47: ROC curves for blue Focus using SAM.



(a) Blue Focus, ASD, SNR=20

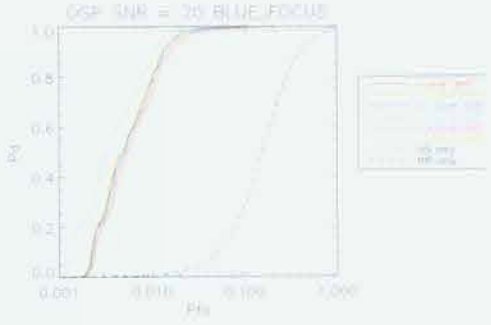


(b) Blue Focus, SNR=100

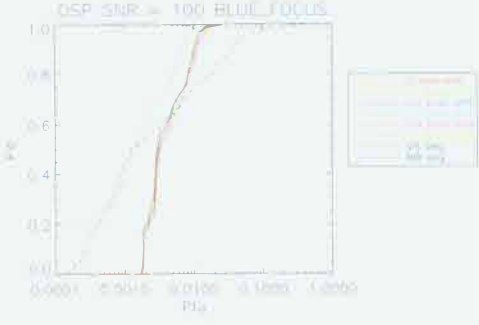


(c) Blue Focus, ASD, SNR=500

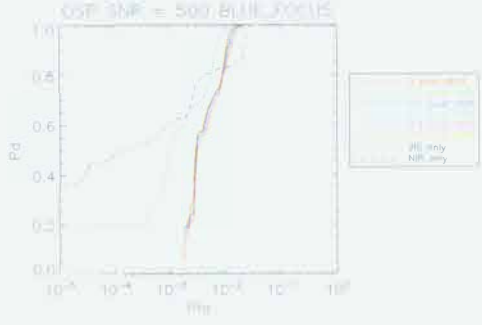
Figure 7.48: ROC curves for blue Focus using ASD.



(a) Blue Focus, OSP, SNR=20

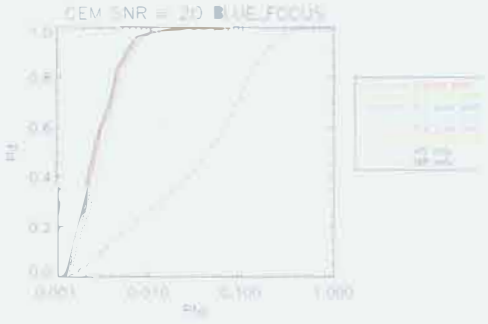


(b) Blue Focus, SNR=100

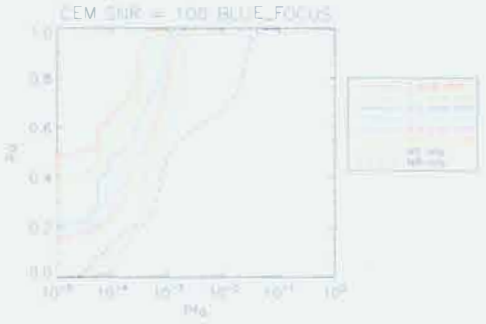


(c) Blue Focus, OSP, SNR=500

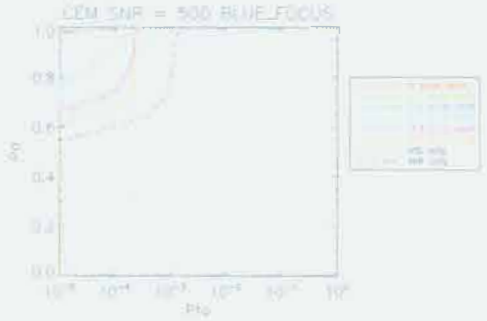
Figure 7.49: ROC curves for blue Focus using OSP.



(a) Blue Focus, CEM, SNR=20

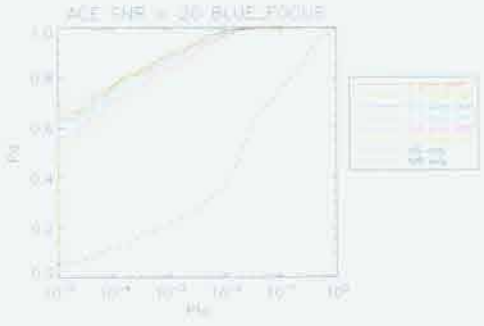


(b) Blue Focus, CEM, SNR=100

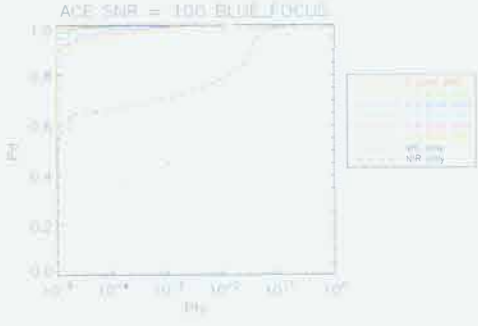


(c) Blue Focus, CEM, SNR=500

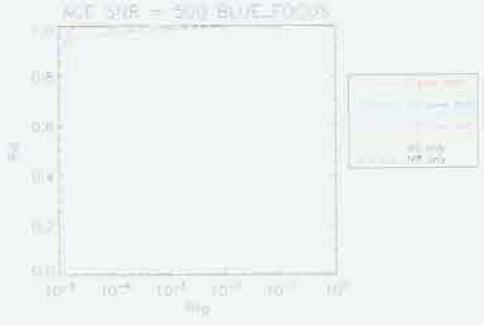
Figure 7.50: ROC curves for blue Focus using CEM.



(a) Blue Focus, ACE, SNR=20

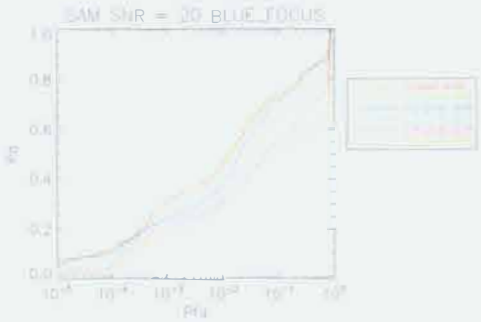


(b) Blue Focus, ACE, SNR=100

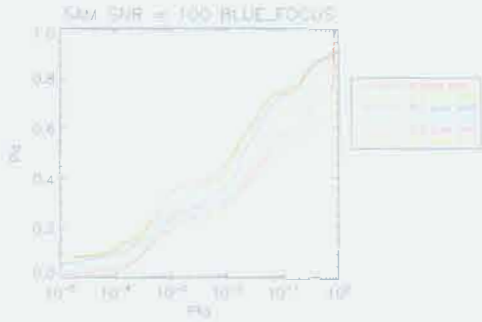


(c) Blue Focus, ACE, SNR=500

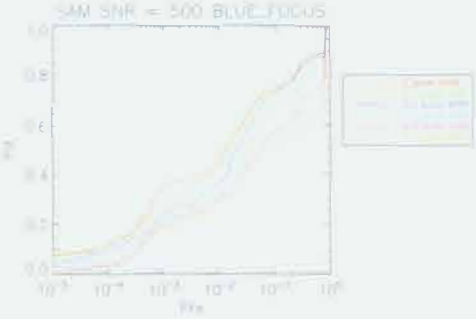
Figure 7.51: ROC curves for blue Focus using ACE.



(a) Blue Focus/grass mix, SAM, SNR=20

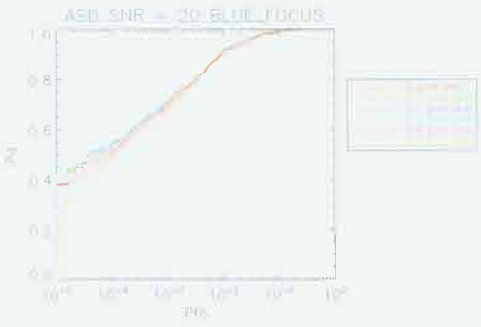


(b) Blue Focus/grass mix, SAM, SNR=100

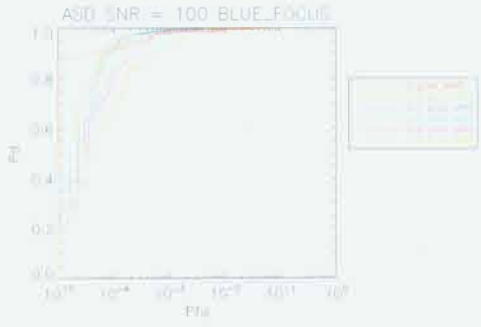


(c) Blue Focus/grass mix, SAM, SNR=500

Figure 7.52: ROC curves for blue Focus/grass linear mixture using SAM.



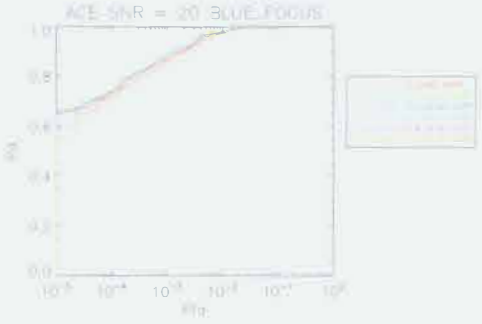
(a) Blue Focus/grass mix, ASD, SNR=20



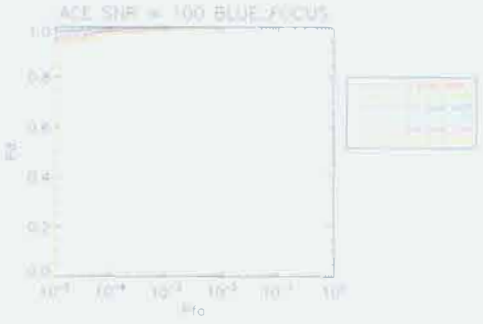
(b) Blue Focus/grass mix, ASD, SNR=100

(c) Blue Focus/grass mix, ASD, SNR=500

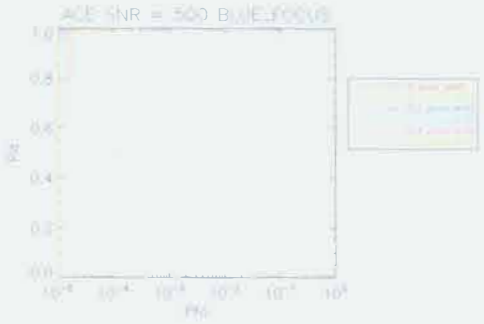
Figure 7.53: ROC curves for blue Focus/grass linear mixture using ASD.



(a) Blue Focus/grass mix, ACE, SNR=20

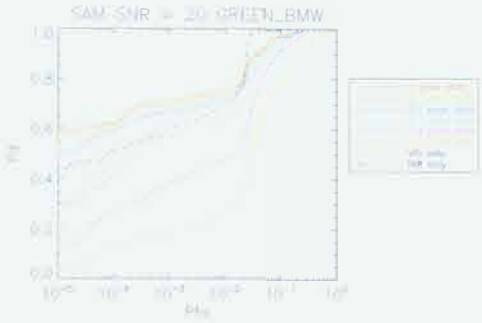


(b) Blue Focus/grass mix, ACE, SNR=100

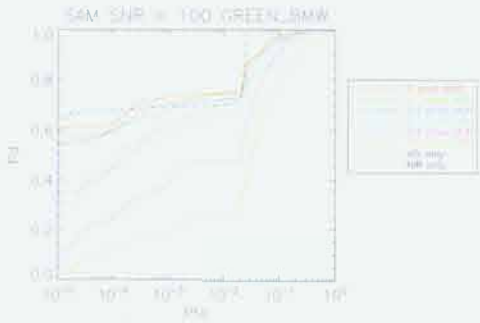


(c) Blue Focus/grass mix, ACE, SNR=500

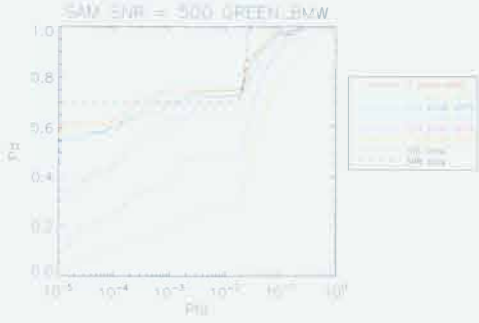
Figure 7.54: ROC curves for blue Focus/grass linear mixture using ACE.



(a) Green BMW, SAM, SNR=20

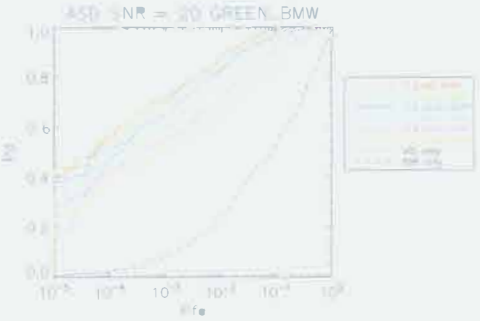


(b) Green BMW, SAM, SNR=100

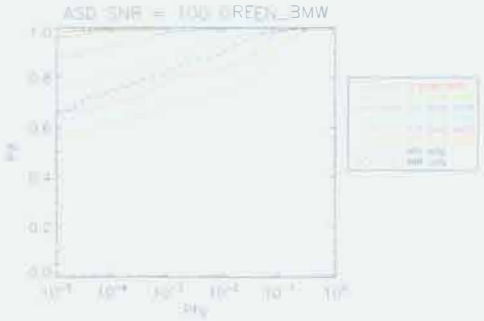


(c) Green BMW, SAM, SNR=500

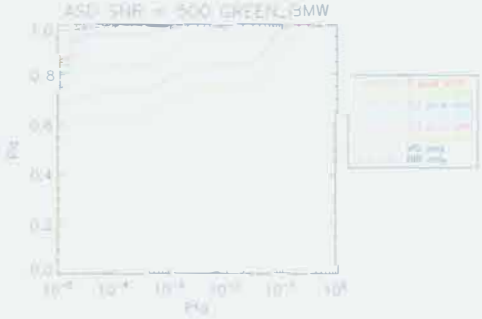
Figure 7.55: ROC curves for green BMW using SAM.



(a) Green BMW, ASD, SNR=20

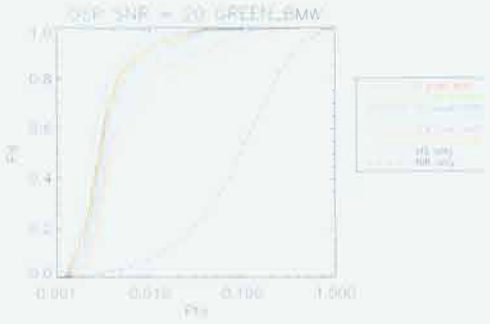


(b) Green BMW, SNR=100

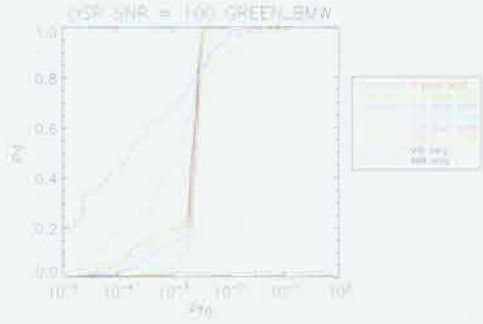


(c) Green BMW, ASD, SNR=500

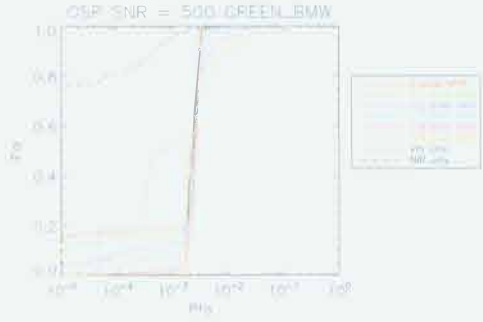
Figure 7.56: ROC curves for green BMW using ASD.



(a) Green BMW, OSP, SNR=20

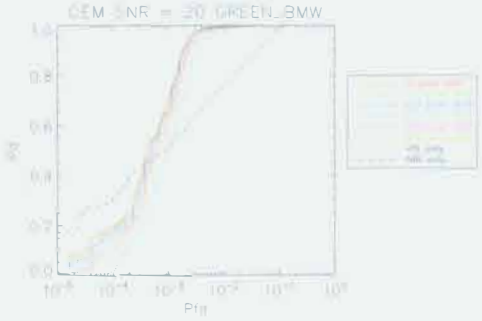


(b) Green BMW, SNR=100

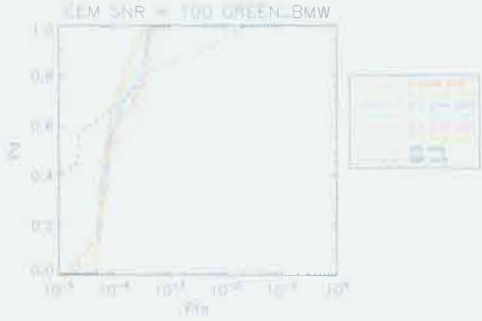


(c) Green BMW, OSP, SNR=500

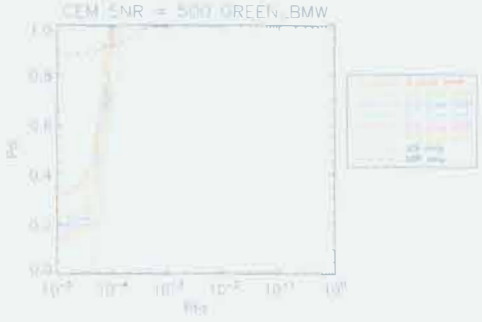
Figure 7.57: ROC curves for green BMW using OSP.



(a) Green BMW, CEM, SNR=20

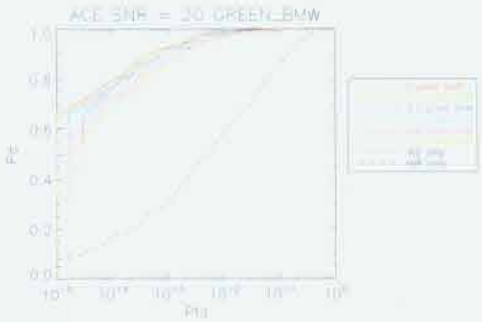


(b) Green BMW, CEM, SNR=100

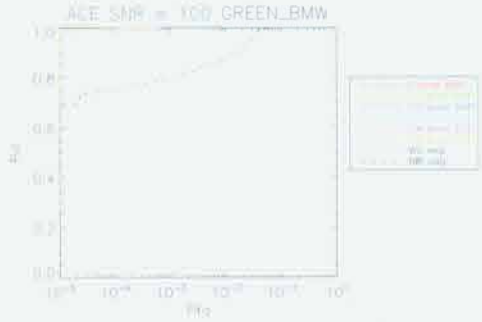


(c) Green BMW, CEM, SNR=500

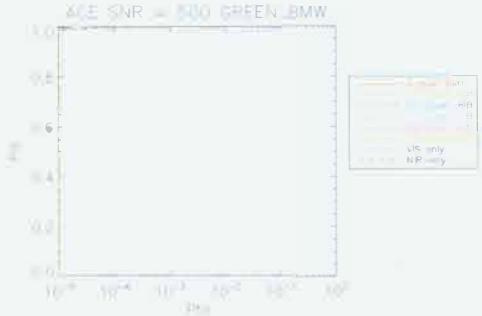
Figure 7.58: ROC curves for green BMW using CEM.



(a) Green BMW, ACE, SNR=20

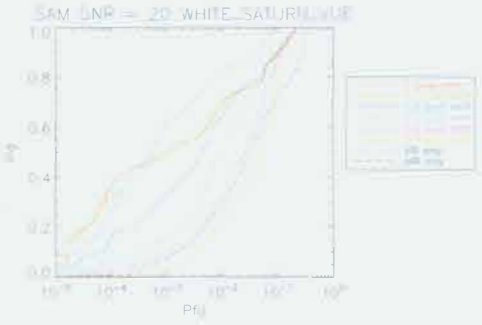


(b) Green BMW, ACE, SNR=100

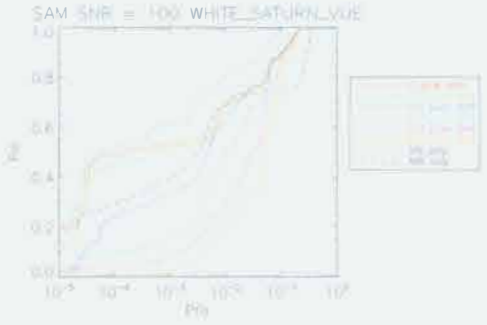


(c) Green BMW, ACE, SNR=500

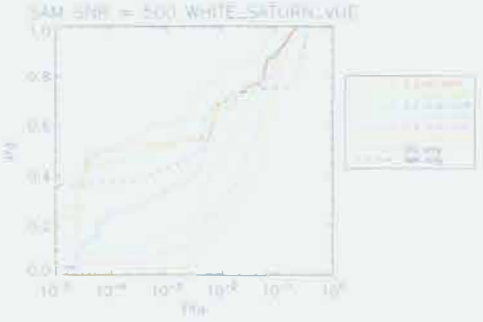
Figure 7.59: ROC curves for green BMW using ACE.



(a) White Saturn Vue, SAM, SNR=20

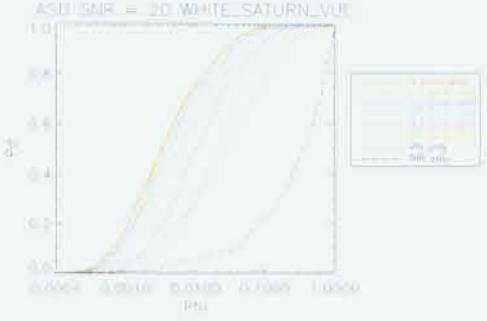


(b) White Saturn Vue, SAM, SNR=100

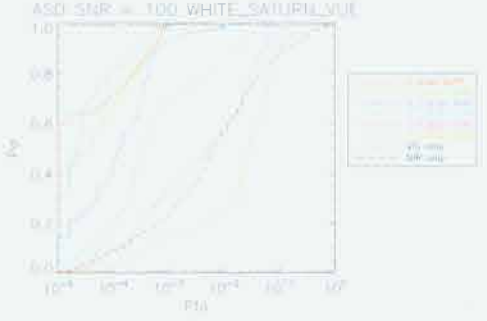


(c) White Saturn Vue, SAM, SNR=500

Figure 7.60: ROC curves for White Saturn Vue using SAM.



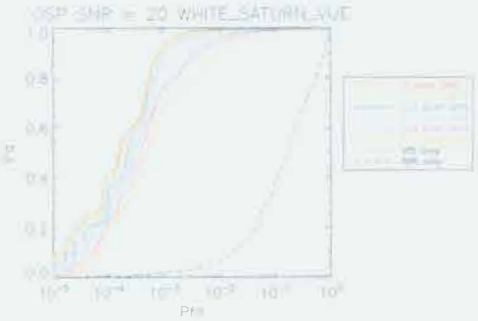
(a) White Saturn Vue, ASD, SNR=20



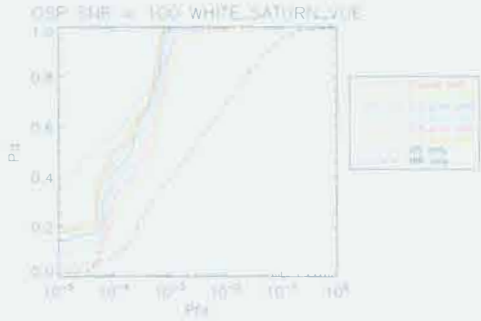
(b) White Saturn Vue, SNR=100

(c) White Saturn Vue, ASD, SNR=500

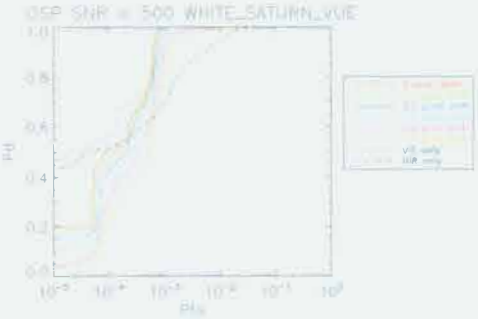
Figure 7.61: ROC curves for White Saturn Vue using ASD.



(a) White Saturn Vue, OSP, SNR=20

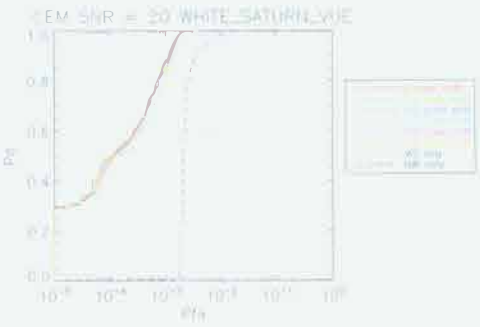


(b) White Saturn Vue, SNR=100

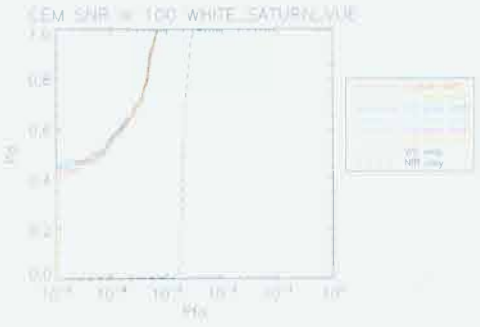


(c) White Saturn Vue, OSP, SNR=500

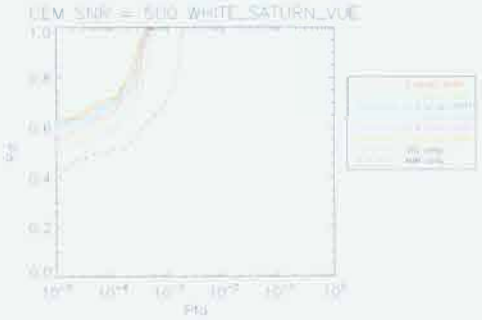
Figure 7.62: ROC curves for White Saturn Vue using OSP.



(a) White Saturn Vue, CEM, SNR=20

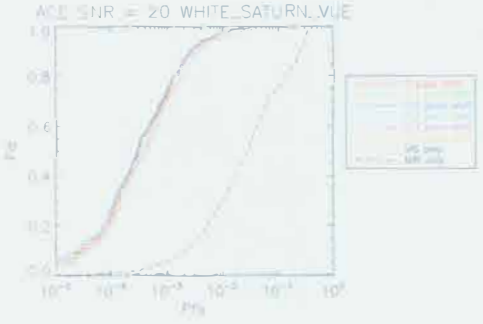


(b) White Saturn Vue, CEM, SNR=100

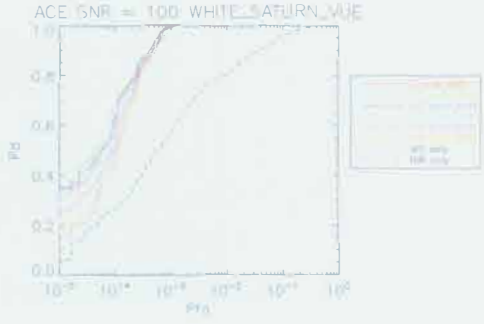


(c) White Saturn Vue, CEM, SNR=500

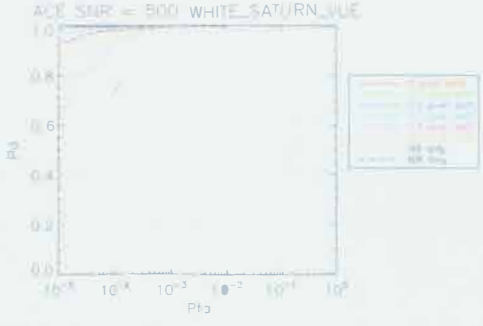
Figure 7.63: ROC curves for White Saturn Vue using CEM.



(a) White Saturn Vue, ACE, SNR=20



(b) White Saturn Vue, ACE, SNR=100



(c) White Saturn Vue, ACE, SNR=500

Figure 7.64: ROC curves for White Saturn Vue using ACE.

7.3 MISI Target Detection Results

Real MISI imagery was also used in an effort to validate this study. The data used was imagery of RIT's campus collected by MISI on June 24th, 2005. The imaged scene is shown in Figure 7.65, and was part of a vehicle tracking experiment. The area inside the red rectangle in Figure 7.65 shows a smaller subset that was used as our scene of interest. The raw MISI imagery was processed by registering the sets of VIS and NIR bands using the methods discussed in chapter 4. The MISI sensor model along with GPS and DMU measurements were used to compute ground location coordinates for each pixel in the VIS and NIR bands. These coordinates were used to perform a nearest neighbor interpolation to register the bands to one another. ELM atmospheric compensation was applied to the scene using calibration panels of known reflectance which were painted in one of the parking lots. Each spectral band was observed in the ELM compensated image, and very noisy bands were removed. The bands that were removed due to noise were the first 5 bands in the VIS and the last 9 bands in the NIR. The methods discussed in section 6.2 were used to characterize the background. The number of endmembers used for k-means was varied in an attempt to optimize target detection performance. Five endmembers were ultimately chosen.

The larger image of the scene of interest is shown in Figure 7.66. This scene contains two vehicles of known reflectance which were used as targets. The two vehicles used in this experiment were a green BMW and a white Saturn Vue, which were also used in the target detection experiment using the DIRSIG scene. The green and white rectangles in Figure 7.66 denote the locations of the green BMW and white Saturn Vue respectively. An enlarged view showing the exact location of the green BMW can be seen in figure 7.67. The green BMW is shown inside the green circle in this figure.

Target detection algorithms SAM, ASD, OSP, CEM and ACE were applied to the scene of interest for the green BMW and white Saturn Vue. The algorithms used were performed using the VIS bands only, as well as the full spectrum. Many of the NIR bands were very noisy and discarded, so the algorithms were not run using only the NIR bands in this case. Detection images for 2 algorithms are shown in Figure 7.68. The detection images shown in Figure 7.68 are of a very small subset of the scene of interest, and the algorithms were run using the full spectrum.

It can be seen from the detection images in Figure 7.68 that this target is not very easily

Figure 7.65: Full view of campus scene captured by MISI

Figure 7.66: Scene of interest for MISI target detection experiment

Figure 7.67: Small subset showing the location of the green BMW

(a) Green BMW detection, MISI imagery, SAM, using full spectrum

(b) Green BMW detection, MISI imagery, ASD, using full spectrum

Figure 7.68: Results using SAM (a) and ASD (b) to detect the green BMW in MISI imagery. All images show roughly the top 5% of pixels.

detected. Similar results were obtained when attempting to detect the white Saturn Vue. There are several reasons why detection performance was rather poor using real MISI imagery. The targets used in this experiment were difficult to detect. The simulated target detection experiments demonstrate that these vehicles were not detected as easily as higher contrast targets such as the red Toyota or blue Ford Focus. MISI suffers from a relatively low SNR, the MISI imagery also had many noisy bands which had to be discarded. The amount of misregistration between spectral bands was also unknown. The weather conditions, including varying illumination conditions due to cloud cover were not known. If illlumination varies over the scene, then the retrieved reflectance for pixels far away from the calibration panels might not be correct. Ground truth was also not perfectly known. The reflectance of the painted panels was recorded when they were first painted onto the parking lot, but weathering, contamination, and vehicles driving over the panels cause the reflectance to change over time. All these factors contribute to the low detection rates of the targets using real imagery. The percentages of all pixels having a higher detection test statistic value than the target are displayed in Table 7.1

Table 7.1: Target Detection performance using MISI data

Green BMW			White Saturn Vue		
VIS only			VIS only		
Algorithm	# FA's	%FA's	Algorithm	#FA's	%FA's
SAM	7033	3.9	SAM	652	0.36
ASD	10007	5.5	ASD	3036	0.17
OSP	6118	3.4	OSP	1379	0.76
CEM	154616	85	CEM	2415	1.3
ACE	748	0.41	ACE	2453	1.4
Full Spectrum			Full Spectrum		
Algorithm	# FA's	%FA's	Algorithm	#FA's	%FA's
SAM	1266	0.70	SAM	49601	27
ASD	20956	12	ASD	1399	0.71
OSP	55313	30	OSP	1014	0.56
CEM	2044	1.1	CEM	2571	1.4
ACE	764	0.42	ACE	2022	1.1

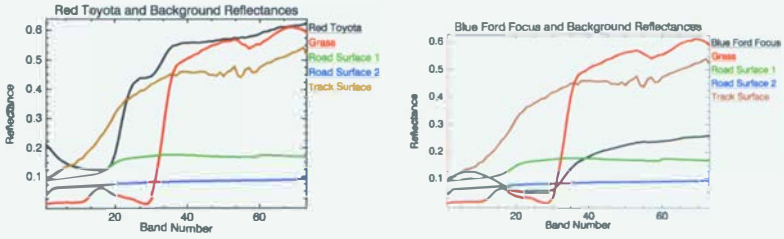
Chapter 8

Discussion and Conclusions

Detection results using the full spectrum can be seen for the red Toyota in Figures 7.19, 7.20, 7.21, 7.22, and 7.23. Using a quick visual inspecton, there are several phenomena apparent in these images. Looking at the SAM results in Figure 7.19, it can clearly be seen that as the level of misregistration increases, many vehicles become more difficult to detect, as we see many false alarms appear in the detection maps. Notice, however, that the vehicles parked in the grass fields are still quite easily detectable, while vehicles parked on the road surfaces tend to disappear for larger misregistration amounts. Looking at the reflectance curves for the various materials in Figure 8.1, we notice that grass, like the red Toyota, has a high reflectance in the NIR. This would account for the reason that red Toyotas are still easily detected in the grass field at high amounts of misregistration. Because spectral mixing between the red Toyota and background is occurring only the the NIR bands, the spectrum isn't disturbed very much for red Toyotas surrounded by grass. The spectra of the red Toyotas surrounded by other background surfaces will change more because the other backgrounds have relatively low reflectance values in the NIR. Also notice the running track surface has a higher reflectance in the NIR than the road surfaces. Red Toyotas parked on the track area are also quite detectable at high levels of misregistration.

This is somewhat easy to see when using a simple algorithm like SAM, which depends only on the angle between the target vector and the pixel of interest. It is not as easy to predict the effects of misregistration on more sophisticated algorithms, so this section will discuss trends seen in the data to attempt to draw meaningful conclusions.

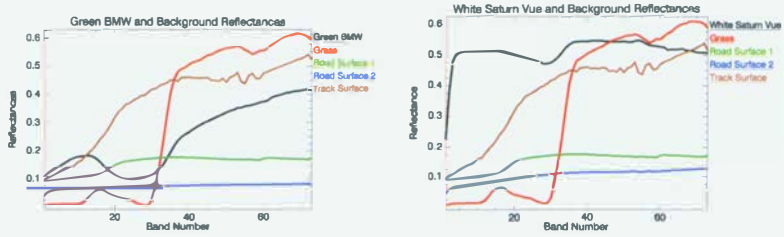
The blue Ford Focus is a much darker target than the red toyota, as can be seen in image (b)



(a) Red Toyota and Background Reflectance

(b) Blue Ford Focus and Background Reflectance

Figure 8.1: Comparison of background, red Toyota and blue Ford Focus reflectances



(a) Green BMW and Background Reflectance

(b) White Saturn Vue and Background Reflectance

Figure 8.2: Comparison of background, green BMW and white Saturn Vue reflectances

of Figure 8.1. The reflectance curve for this target is much more similar to the darker background materials, mainly the two road surfaces. As the amount of misregistration increases, again, the detections maps show the blue Ford Focuses parked against the dark road surfaces are more easily detected because the target pixel spectrum is not altered as much as the target pixels in the grass and track backgrounds. This effect is greatly noticeable in Figure 7.25, the resulting detection maps from ASD. At registration levels of 0.4 pixel the detections for vehicles parked in grass are very faint, and the vehicles parked on the track are not detected at all. Other target detection algorithms show similar results, even Figure 7.28, the detection maps when using ACE show a lower response for vehicles in grass and on the track. It is interesting to note, however, that detections using the stochastic algorithms tend to give better detection results for this vehicle, and are more robust than geometric algorithms, especially at high levels of misregistration.

In chapter 7, a linear mixing model was used in only the NIR set of bands in an attempt to better detect blue Ford Focus targets in the grass background. Referring to detection maps in Figures 7.29 and 7.30 we see that the vehicles in the grass background stand out more than

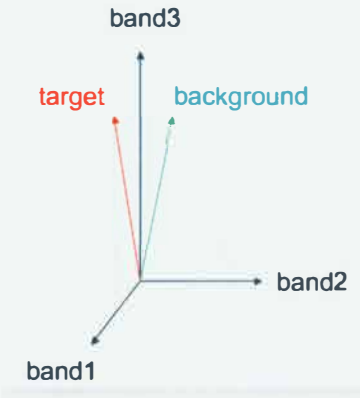


Figure 8.3: 3 band representation of a background and target vector.

those in the detection maps in Figures 7.24 and 7.25. We also notice that more false alarms appear on the edges, and the vehicles parked on other background surfaces do not stand out as much. Detectability appears to go down overall, which is clearly demonstrated when comparing ROC curves in Figures 7.47 and 7.48 to those in Figures 7.52 and 7.53. Here, using a linear mixing model in one set of bands does improve detection for targets in a particular background, but is not a viable solution if the surrounding material is unknown.

The ROC curves created for each algorithm run for each level of misregistration can be seen in chapter 7. Each curve was plotted in the same window for each algorithm to easily compare detection for each level of misregistration at any given false alarm rate. Several phenomena appear to be somewhat consistent in the results. SAM and ASD tend not to benefit from using the full spectrum. This is surprisingly true even when there is very little or no misregistration in some cases, especially when applying these algorithms to imagery with higher SNRs. This may at first seem counter-intuitive as adding spectral bands would provide additional useful data for classifying pixels as target or background. A simple 3 band example is constructed seen in Figure 8.3, which demonstrates that adding spectral bands is not always useful. Recalling equation 3.5, which is the equation for SAM, target and background pixels may be compared by computing the angle between the two vectors. Here, the target vector is given in equation 8.1, and the background vector is given in equation 8.2.

$$target = [1, 0, 3] \tag{8.1}$$

$$background = [0, 1, 3] \tag{8.2}$$

Using only the first 2 bands of the target and background vectors, the angle between the two vectors is 90 *degrees*. Using all 3 bands, the angle between the background and target vectors is ≈ 26 *degrees* which means this particular background vector appears more target-like using 3 bands as opposed to just the first 2 bands. This demonstrates, at least using SAM, that adding spectral bands does not necessarily improve detection performance, and in some cases, adding spectral bands may be counter-productive.

It can be seen from Figures 7.11, 7.12, 7.13, and 7.14 that using only the VIS bands usually results in better detection as opposed to using only the NIR bands. This is true for both detecting both the blue Focus and the red Toyota. It is more difficult to tell if this is true for the green BMW and white Saturn Vue. It appears in some cases using only the NIR bands results in better detection performance for these vehicles. The ROC curves show that the performance using the VIS bands only and NIR bands only depends on many factors. It seems to change dramatically based on the vehicle, the SNR and the chosen false alarm rate.

The green BMW is a target which is a bit more difficult to analyze than the red Toyota or blue Ford Focus. Image (a) in Figure 8.2 compares the spectrum of the green BMW and the various background spectra. The green BMW has a unique spectra in the VIS region of the spectrum (bands 1-35) as well as the NIR (bands 36-73). Notice the target reflectance in the NIR is closer to the road surfaces than grass for the lower bands in the NIR (bands 35-45). The target reflectance becomes closer to the track surface and grass for the higher bands in the NIR. The target spectrum will be disturbed no matter what background this target is parked against, but the spectrum will be affected in different regions depending on the background spectrum.

The white Saturn Vue is also a difficult target to analyze for several reasons. One reason is this target is very spectrally flat. This makes this target difficult to detect, especially when using an algorithm like SAM, as it is spectrally uninteresting, and not likely a very unique signature in the scene. Referring to Figure 7.17, we can see the detection maps for the white Saturn Vues using the SAM, ASD and OSP detectors run on the VIS and NIR channels separately. The vehicles parked against the track and on the grass have fewer returns than the vehicles parked against the other background surfaces, especially when looking at the detection maps for SAM and ASD. As misregistration increases, these targets will become even more difficult to detect.

On the other hand, the white Saturn Vue, although spectrally flat in the NIR, has reflectance values closer to the track and grass. As spectral mixing between background and target pixels

occurs for vehicles parked in grass and track backgrounds, the overall test pixel spectrum is not altered as much as test pixels in a darker background. Since the resulting spectrum is not altered as much for vehicles parked against bright backgrounds, we expect these vehicles will be easier to detect at higher misregistration amounts than vehicles parked against darker backgrounds. Each algorithm seems to produce slightly different results. Performance tends to drop off steeply as misregistration increases when using SAM and ASD, as can be seen in Figures 7.37 and 7.38. For these algorithms, performance seems to drop off for every vehicle, regardless of background. Figures 7.39, 7.40 and 7.41 show results that indicate performance does not degrade that much as misregistration increases. These results are pretty consistent as far as the stochastic algorithms are concerned, but it is somewhat surprising to see the robustness of OSP for this target.

A few trends appear to remain somewhat consistent when looking at all the ROC curves in Figures 7.42 through 7.64. First, target detection performance tends to be greatly affected by band-to-band misregistration. For each target, there is quite a large discrepancy between the PD with no misregistration and 0.5 pixel misregistration for any given FAR. Some targets benefit from applying SAM to images utilizing the full VNIR spectrum, while some targets are detected more easily using only the VIS bands. ASD, while appearing more robust than SAM, still suffers from large performance degradations as misregistration increases. This can be difficult to see for some targets, such as the red Toyota, which overall is an easy target to detect. Observing some of the more difficult targets, especially looking at the ROC curves for the white Saturn Vue in Figure 7.61, the disparity between the performance at 0 pixel misregistration and 0.5 misregistration is quite large at any FAR. OSP, CEM and ACE appear to be much more robust in the presence of misregistered data. Many curves appear to be very similar, almost identical even at 0.5 pixel misregistration. One major difference between the OSP algorithm and the stochastic algorithms, is that OSP, although a robust target detector when dealing with misregistered images, does not seem to benefit from the addition of NIR bands, even at small or zero misregistration levels. CEM and ACE, on the other hand, benefit from the addition of the NIR bands, even at very large misregistration amounts. It is also interesting to note how well ACE performs in each case, especially at higher signal to noise ratios. In some cases, in fact, perfect detection is achieved at any level of misregistration.

The results using synthetic data show statistical algorithms tend to be more robust than geometric algorithms (including SAM) when applying target detection algorithms to images

with band-to-band misregistration. ACE in particular shows very strong potential for use in imagery with uncertainties in the amount of spatial misregistration between sets of spectral bands. These conclusions have been derived from observing the detection maps, ROC curves for each detection image, and are shown side by side as a set of bar graphs in Figures 8.4, 8.5, 8.6 and 8.7. The results are given using images with 2 different SNR's in each case. In Figure 8.4, the results using an SNR of 100 are uninteresting because the red Toyotas are easy to detect and each algorithm shows good performance. Using SNR of 20 for the image containing red toyotas, detections go down slightly for the geometric algorithms, but the performance using the statistical algorithms remains very strong. Figure 8.5 also demonstrate the robustness of ACE. Target detection performance is relatively poor for most of the other algorithms at a SNR of 100. CEM gets a large performance gain when increasing the SNR to 500 while the geometric algorithms appear to have only minimal gains. We see similar results in Figures 8.6 and 8.7, with ACE often outperforming the other algorithms.

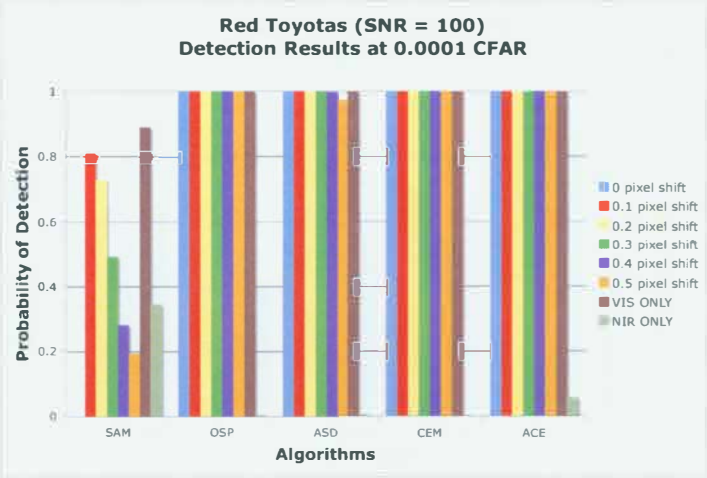
The effects misregistration have on target detection performance have been demonstrated in the previous chapters, but there is room for future work. Future work should include another data collection with MISI or another hyperspectral sensor suffering band-to-band misregistration. As seen chapter 7, the real data being used was taken with an imaging system with a low SNR, the atmospheric and illumination conditions were unknown, there was questionable ground truth, and low contrast or spectrally flat targets were used. These factors make the target detection process using real imagery very difficult. Another experiment under clear and stable atmospheric conditions with reliable ground truth and spectrally unique, high contrast targets would be useful for validation of the simulated experiments as well as further research.

A method using spectral mixing in only one set of spectral bands was used for the blue Ford Focus in the simulated experiments. This method used the target spectrum and background spectrum for only one background material. More research may be conducted to develop new target spaces which vary the amount of misregistration, as well as the background material the target is surrounded by. By employing these methods new algorithms could be developed which are invariant to the amount of band-to-band misregistration as well as the background surrounding the target.

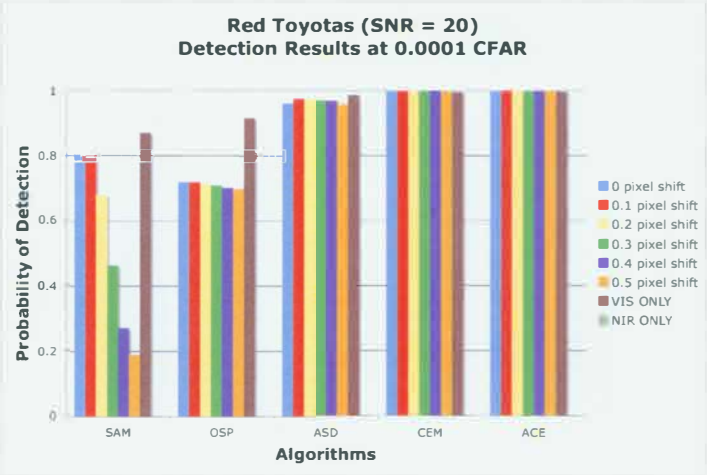
Some of the ROC curves shown in Chapter 7 produce minimal or non-intuitive changes at various signal-to-noise ratios. For example, Figure 7.43 shows decreasing detection performance as SNR increases. Further investigation is required to determine why this occurs. The general

trends in the results show stochastic algorithms performing better overall at large amounts of misregistration, while Geometric algorithms tend to perform well with no or very little misregistration. ACE tends to perform well for almost every tested scenario, even at high levels of misregistration.

Several recommendations on processing misregistered data for target detection can be made from the trends observed in this work. Since target detection performance degrades as misregistration between bands increases, spectral bands should be registered to one another as precisely as possible using a sensor model when using a hyperspectral sensor which suffers from band-to-band misregistration. Geometric algorithms perform well using only 1 set of spectral bands, but often perform very poorly when using the other set of bands. It is also difficult to determine which set of bands will give better performance. The full spectrum should be used for target detection applications because of this uncertainty. Stochastic algorithms, and ACE in particular show good performance compared to geometric algorithms in this case. It is recommended that ACE be the target detection algorithm of choice when performing spectral target detection on misregistered hyperspectral images.

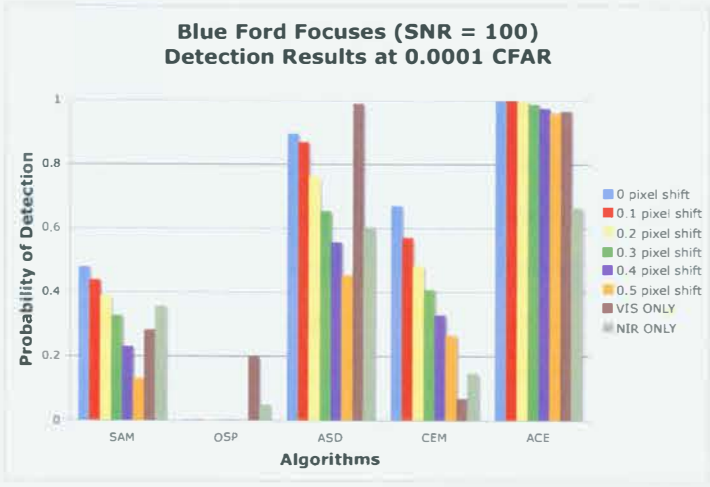


(a) Red Toyota Detection Results at a CFAR for SNR:100

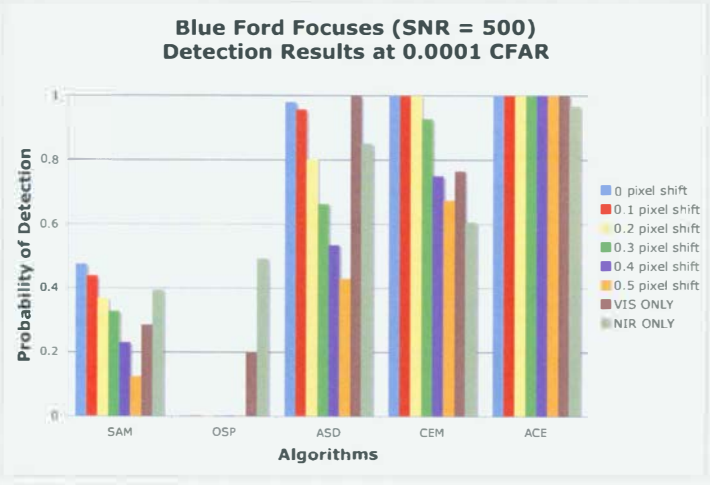


(b) Red Toyota Detection Results at a CFAR for SNR:20

Figure 8.4: Detection results for the red Toyotas with an SNR of 100(a) and SNR of 20 (b) at a CFAR of 0.0001. Detections for each algorithm and each amount of misregistration are shown here.

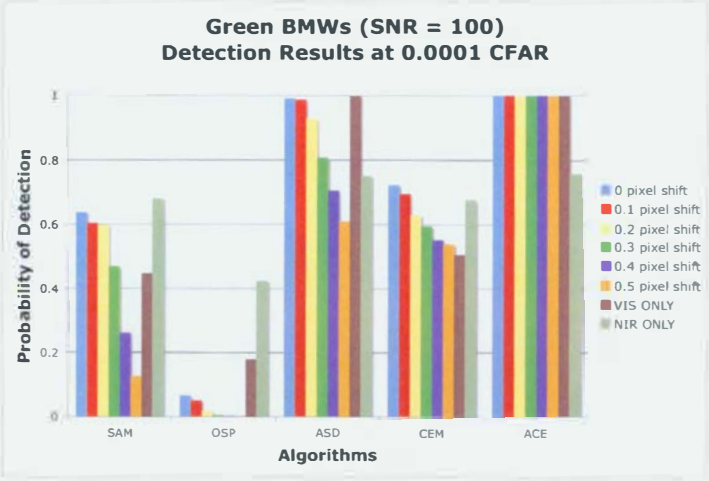


(a) Blue Ford Focus Detection Results at a CFAR for SNR:100

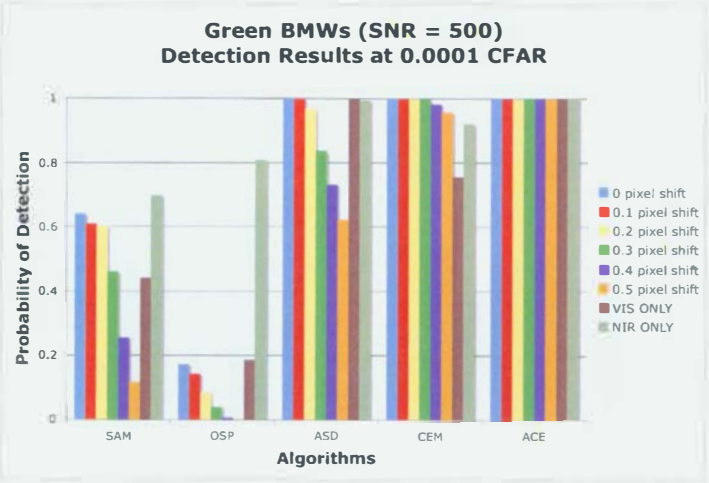


(b) Blue Ford Focus Detection Results at a CFAR for SNR:500

Figure 8.5: Detection results for the red Toyotas with an SNR of 100(a) and SNR of 500 (b) at a CFAR of 0.0001. Detections for each algorithm and each amount of misregistration are shown here.

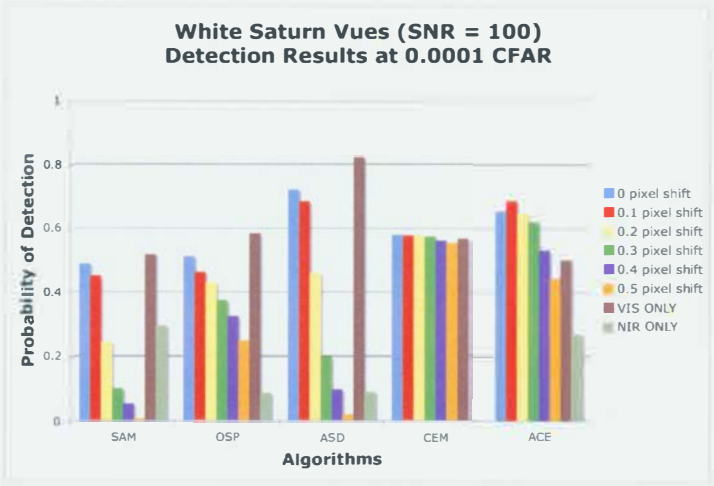


(a) Green BMW Detection Results at a CFAR for SNR:100

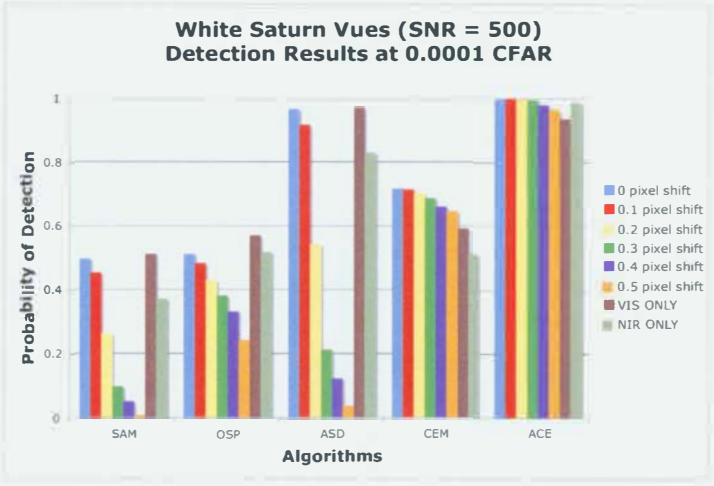


(b) Green BMW Detection Results at a CFAR for SNR:500

Figure 8.6: Detection results for the green BMWs with an SNR of 100(a) and SNR of 500 (b) at a CFAR of 0.0001. Detections for each algorithm and each amount of misregistration are shown here.



(a) White Saturn Vue Detection Results at a CFAR for SNR:100



(b) White Saturn Vue Detection Results at a CFAR

Figure 8.7: Detection results for the white Saturn Vues with an SNR of 100(a) and SNR of 5●● (b) at a CFAR of ●.●●●1. Detections for each algorithm and each amount of misregistration are shown here.

Bibliography

[1] T. Reichardt, "The first photo from space," *Air and Space Magazine*, October-November 2006. 1

[2] *The DIRSIG User's Manual*, Digital Imaging and Remote Sensing Laboratory, Chester F. Carlson Center for Imaging Science Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, 14623, July 2006. [Online]. Available: http://dirsig.cis.rit.edu/doc/manual/ 1.1, 1.2, 1.3

[3] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The hymap airborne hyperspectral sensor: The system, calibration and performance," in *EARSEL Workshop on Imaging Spectroscopy*, Zurich, October 1998. 1

[4] R. J. Birk and T. B. McCord, "Airborne hyperspectral sensor systems," *IEEE AES Systems Magazine*, pp. 26–33, October 1994. 1

[5] J. Schott, S. Brown, R. Raqueño, G. Gross, and G. Robinson, "An advanced synthetic image generation model and its application to multi/hyperspectral algorithm development," *Canadian Journal of Remote Sensing*, vol. 25, no. 2, June 1999. 2

[6] J. R. Schott, *Remote Sensing, The Image Chain Approach.* Oxford University Press, 1997, no. pg. 114. 3.2.1

[7] A. Berk, G. Anderson, P. Acharya, J. Chetwynd, L. Bernstein, E. Shettle, M. Matthew, and S. Adler-Golden, *MODTRAN4 User's Manual*, Air Force Research Laboratory Space Vehicles Directorate, Hanscom AFB, MA, 01731-3010, June 1999. 3.2.1.2

[8] A. Berk, S. Adler-Golden, A. Ratkowski, G. Felde, G. Anderson, M. Hoke, T. Cooley, J. Chetwynd, J. Gardner, M. Matthew, L. Bernstein, P. Acharya, D. Miller, and P. Lewis,

"Exploiting modtran radiation transport for atmospheric correction: The flaash algorithm," in *Fifth International Conference on Information Fusion*, Annapolis, 2002. 3.2.1.2

[9] R. A. Johnson and D. W. Wichern, *Applied Multivarite Statistical Analysis.* Prentice Hall, 2002, vol. 5. 3.2.2.1, 3.3.3.1

[10] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, January 1988. 3.2.2.2

[11] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. Heidebrecht, A. T. Shapiro, P. Barloon, and A. Goetz, "The spectral image processing system (sips) interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of the Environment*, vol. 44, pp. 145–163, 1993. 3.3.1

[12] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: anorthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 779–785, 1994. 3.3.2.1

[13] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, p. 29, January 2002. 3.3.2.2

[14] E. Conte, M. Lops, and G. Ricci, "Asumptotically optimum radar detection in compound-gaussian clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, pp. 617–625, 1995. 3.3.3.1

[15] D. Manolakis, G. Shaw, and N. Keshava, "Comparative analysis of hyperspectral adaptive matched filter detectors," *SPIE Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, vol. 4049, pp. 2–17, 2000. 3.3.3.1, 3.3.3.1

[16] D. Manolakis, C. Siracusa, D. Marden, and G. Shaw, "Hyperspectral adaptive matched filter detectors: Practical performance comparison," in *Algorithms for Multispectral and Hyperspectral Imagery*, S. S. Shen and M. R. Descour, Eds., vol. 4381. Orlando, FL: Proceedings of SPIE, 2001. 3.3.3.1

[17] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions,"

IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 6, pp. 2706–2717, November 1999. 3.4.1

[18] H. Akima, "A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points," *ACM Transactions on Mathematical Software (TOMS)*, vol. 4, pp. 148–159, June 1978. 3.5

[19] C. Lawson, "Generation of a triangular grid with application to countour plotting," Jet Propulsion Lab, California Institute of Technology, Pasadena, California, Tech. Rep. 299, February 1972. 3.5

[20] R. Wiemker, "Registration of airborne scanner imagery using akima local quintic polynomial interpolation," in *Proceedings of the Second International Airborne Remote Sensing Conference and Exhibition*, vol. 3, San Francisco, Ann Arbor, 1996, pp. 210–219. 3.5

[21] W. Zhang and J. Albertz, "Rectification of airborne line-scanner imagery utilizing flight parameters," in *Proceedings of the First International Airborne Remote Sensing Conference and Exhibition*, vol. 2, Strasbourg, France, 1994. 3.5

[22] J. P. Kerekes, M. Muldowney, K. Strackerjan, L. Smith, and B. Leahy, "Vehicle tracking with multi-temporal hyperspectral imagery," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII*, vol. 6233, Orlando, FL, April 2006. 3.6

[23] E. Hecht, *Optics*, 4th ed. Addison Wesley, August 2001. 4

[24] J. Faulring, "private communication," CIS RIT, Fall 2007. 4

[25] S. R. Lach and D. McKeown, "private communication," CIS RIT, Fall 2006. 4

[26] P. R. Wolf and B. A. Dewitt, *Elements of Photogrammetry with Applications in GIS*, 3rd ed. McGraw hill, February 2000. 4

[27] P. H. Swain, V. C. Vanderbilt, and C. D. Jobusch, "A quantitative applications-oriented evaluation of thematic mapper design specifications," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-20, no. 3, pp. 370–377, July 1982. 6

[28] P. Mouroulis, D. A. Thomas, T. G. Chrien, V. Duval, R. ●. Green, J. J. Simmonds, and A. H. Vaughan, "Trade studies in multi/hyperspectral imaging systems," Jet Propulsion Laboratory, Tech. Rep., 1998. 6

[29] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1566–1577, September 1998. 6

[30] E. J. Ientilucci and S. D. Brown, "Advances in wide area hyperspectral image simulation," in *Targets and Backgrounds IX: Characterization and Representation*, W. R. Watkins, D. Clement, and W. R. Reynolds, Eds., vol. 5075, SPIE. Proceedings of SPIE, 2003. 6.1.1

[31] D. Elliot, L. Hung, E. Leong, and D. Webb, "Accuracy of gaussian approximation for simulating eo sensor response," *Signals, Systems and Computers, 1996. 1996 Conference Record of the 30th Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 868–872, November 1996. 6.1.2